7-2020

Taxonomic and Genetic Diversity, and Pathogenicity of Diaporthe Species Associated with Soybean

Fakhir Eraheem Hameed Al Shuwaili
University of Arkansas, Fayetteville

Taxonomic and Genetic Diversity, and Pathogenicity of *Diaporthe* Species Associated with Soybean

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Cell and Molecular Biology

by

Fakhir Eraheem Hameed Al Shuwaili
University of Baghdad
Bachelor of Science in Plant Protection, 1999
University of Baghdad
Master of Science in Plant Protection, 2002

July 2020
University of Arkansas

The dissertation is approved for recommendation to the Graduate Council.

_____
Burt H. Bluhm, PhD
Dissertation Director

_____
John C. Rupe, PhD
Committee Member

_____
Mary C. Savin, PhD
Committee Member

_____
Alejandro Rojas, PhD
Committee Member

Abstract

Diaporthe species (anamorph: *Phomopsis*) are associated with a wide range of plant hosts as plant pathogens, asymptomatic endophytes, and saprobes. One of these hosts is soybean, which is one of the most important crops in U.S. agriculture. Several *Diaporthe* species cause important diseases on soybean in the U.S., and specifically in Arkansas. The taxonomy, genetic diversity, and pathogenicity of *Diaporth*e species associated with asymptomatic infection of soybean are rarely studied with accurate molecular tools. Therefore, this dissertation aimed to assess the diversity and boundaries of *Diaporth*e associated with soybean in Arkansas. Furthermore, pathogenicity and alternative lifestyles were assessed among *Diaporthe* strains originating from Arkansas. Moreover, the molecular basis of pathogenesis was dissected in the most ubiquitous *Diaporthe* species, *D. longicolla,* via forward genetic screening. Phylogenetic analyses of multilocus data identified two pathogenic *Diaporthe* species in Arkansas besides the most common species, *D. longicolla.* In this study, *D. unshiuensis* was reported for the first time in the U.S. and on soybean worldwide, while *D. ueckerae* was recorded in Arkansas for the first time. Pathogenicity tests confirmed that these species could potentially alternate between endophytic and pathogenic lifestyles on soybean. Genotyping by sequencing (GBS) utilizing microsatellites revealed that *D. longicolla and D. unshiuensis* had high levels of genetic variability at all study sites. Additionally, these markers successfully discriminated isolates of *D. longicolla, D. ueckerae,* and *D. unshiuensis*. Genotypes of these species did not cluster genetically based on geographical origin. However, *D. ueckerae,* and *D. unshiuensis* were not isolated from all sites sampled. According to linkage disequilibrium indices, populations of *D. unshiuensis* may undergoing sexual reproduction and random mating, whereas populations of *D. longicolla* may be largely clonal in Arkansas. Furthermore, genetic screening to identify

pathogenicity genes of *D. longicolla* highlighted the potential role of a putative cytochrome p450 in seed colonization, stem necrosis, and asexual reproduction. Together, these findings will help inform the development of new strategies to manage soybean diseases caused by *Diaporthe* species and to augment host resistance.

Table of Contents

Chapter 1: General Introduction

Overview

Soybean, *Glycine max* (L.) Merr, is one of the most important crops grown in the U.S., and Arkansas is a leading soybean producer among southern states. In 2016, 3.1 million acres of soybean were grown in Arkansas, which produced about 145.7 million bushels of harvested seeds valued at approximately $1.44 billion USD (USDA-NASS, 2017).

Plant diseases cause significant soybean yield losses in the U.S. Soybean diseases were estimated to incur economic losses of $81.39 billion USD in the U.S. from 1996 to 2016 (Bandara et al., 2019). Some of the most important pathogens are *Diaporthe/Phomopsis* species which cause *Phomopsis* seed decay, pod and stem blight, and stem canker. Estimates of soybean yield reductions in 2018 caused by these diseases in the U.S. were approximately 22.75, 5.11, and 2.34 million bushels, respectively (Allen et al., 2019). Estimated monetary losses according to marketing year average price per bushel ($8.93 USD in 2018) are approximately $203.16, $45.63, and $20.89 million USD, respectively.

The genus *Diaporthe*

Diaporthe Nitschke (1870) is a fungal genus that belongs to the family Diaporthaceae, order Diaporthales, subclass Diaporthomycetidae, class Sordariomycetes, subphylum Pezizomycotina, and phylum Ascomycota. The genus contains 931 recorded species in Mycobank, (accessed May 2019), while 950 recorded species were listed in Mycobank for *Phomopsis* (Sacc.) Sacc (1905), which is considered the asexual stage of *Diaporthe*. Since *Diaporthe* was named before *Phomopsis, Phomopsis* species should be treated as *Diaporthe* species based on priority (Rossman et al., 2015). *Diaporthe* species were originally recorded based on host association, were subsequently revised based strictly on morphological characters

(Wehmeyer, 1933), and were most recently organized based on phylogenetic relationships and morphological features (Gomes et al., 2013).

Diaporthe species cause devastating plant diseases of various economically important crops (Gomes et al., 2013; Thompson et al., 2015; Udayanga et al., 2015), live as endophytes within a wide range of plant hosts (Gomes et al., 2013), cause human (Garcia-Reyne et al., 2011) and animal diseases (Williamson et al., 1994), function as saprophytes to decompose organic matter (Udayanga et al., 2011), and bioremediate environmental waste (Ting et al., 2016). Furthermore, *Diaporthe* species have also been explored as producers of economically important enzymes and secondary metabolites (Dai et al., 2005; Elsäesser et al., 2005; Isaka et al., 2001; Kobayashi et al., 2003). Some of these products could have antibiotic (Bandre & Šašek, 1977; Dettrakul et al., 2003; Lin et al., 2005), anticancer (Kumaran & Hur, 2009), or bioherbicidal activity (Ash et al., 2010).

Soybean diseases caused by *Diaporthe/Phomopsis* species

Several *Diaporthe/Phomopsis* species have been associated with soybean diseases. Specifically, *P. longicolla* Hobb (syn. *D. longicolla* (Hobb) Santos, Vrandečić & A.J.L. Phillips) causes Phomopsis seed decay; *D. caulivora* (Athow & Caldwell) J.M. Santos, Vrandečić & A.J.L. Phillips causes northern stem canker; *D. aspalathi* Jansen, Castl. & Crous (formerly *D. phaseolorum* var. *meridionalis*) causes southern stem canker, and *D. phaseolorum* var. *sojae* (S. G. Lehman) Wehmeyer causes pod and stem blight (Santos et al., 2011).

Phomopsis seed decay (PSD)

Phomopsis seed decay (PSD) of soybean is widespread and can cause substantial yield losses (Allen et al., 2019). PSD is ubiquitous throughout soybean-growing areas of the U.S. and

other regions of the world (Sinclair, 1993). PSD has become a serious production problem in recent decades throughout the mid-southern U.S. This is, in part, due to the adoption of the Early Season Production System (ESPS). In Arkansas and other mid-southern states, the ESPS is utilized to avoid late-season moisture deficits. Early-maturing cultivars, e.g. MG III and IV, are planted in April or May (Mayhew & Caviness, 1994). While the ESPS can avoid late-season droughts, hot, humid conditions during seed filling and maturation are favorable for PSD development. Therefore, PSD has increased in incidence and severity in these regions (Gillen et al., 2012).

Symptoms of PSD include shriveled, elongated, or cracked seeds, often with a chalky-white appearance. PSD reduces soybean seed quality by decreasing oil content, changing seed composition, and increasing the presence of other molds (Li, 2011). Seed infection can also cause pre- and post-emergence damping-off (Kulik et al.,1999a). Infection is thought to occur via soybean pods, at which point the pathogen invades the ovule and developing seeds through the funiculus and hilum. After *P. longicolla* directly penetrates an immature pod, the pathogen could spread to immature seeds and directly penetrate seed coats. Then, it could colonize the entire seed coat and cotyledons. However, the pathogen has not been observed to penetrate through natural openings of the pod surface (Baker et al., 1987).

PSD is primarily associated with *D. longicolla*, although additional *Diaporthe*/*Phomopsis* spp. may also cause PSD. *D. longicolla* is the predominant species associated with stems, pods, and seeds of soybean, compared with *D. phaseolorum* var. *caulivora* and *D. phaseolorum* var. *sojae* (Xue et al., 2007). However, under irrigation, recovery of *D. longicolla* from leaves, stems, and pods was much higher than from roots. Additionally, seed infection in the same study correlated with pod infection of soybean plants in various environments (Mengistu et al., 2009).

Pod and stem blight

Pod and stem blight of soybean was reported in 1920 for the first time in North Carolina as *Phoma* blight caused by *Phoma* sp. (Wolf & Lehman, 1920). The causal agent was later renamed *Diaporthe sojae* Lehman, but a full, formal species description was not provided (Lehman, 1923). The fungus was later renamed as *D. phaseolorum* var. *sojae* (Lehman) Wehm after a more complete description of asexual and sexual stages was recorded (Wehmeyer, 1933). Two decades later, *D. phaseolorum* var. *batatatis* (previously reported as a pathogen of sweet potato) was reported to cause girdling stem cankers on soybean. However, the pathogen associated with soybean was later recognized as a distinct (and novel) variety*, D. phaseolorum* var*. caulivora* (Athow & Caldwell, 1954). Decades later, an undescribed *Phomopsis* species, which produced only pycnidia in culture and could be morphologically distinguished from *D. phaseolorum* var. *sojae* and *D. phaseolorum* var. *caulivora*, was also associated with pod and stem blight (Kmetz, 1975). This fungus was subsequently described as *Phomopsis longicolla* Hobbs (Hobbs et al., 1985). The formation of linear rows of dark pycnidia on dead or senescing stems, pods, and petioles are distinct symptoms of this pathogen (Kulik & Sinclair, 1999b).

Stem canker

Some *Diaporthe*/*Phomopsis* species cause stem canker symptoms upon infection of soybean. Soybean stem canker was initially described in Iowa, and the causal agent was first identified as *D. phaseolorum* var. *batatatis,* followed by *D. phaseolorum* var. *caulivora* (Athow & Caldwell, 1954). Symptoms usually initiate at nodes as brick-red lesions that become darker, elongated, and sunken. Lesions often girdle soybean stems, which results in dead shoots that retain dead leaves (Backman, 1985). Two distinct stem canker diseases have been proposed on soybean: northern stem canker and southern stem canker (Hobbs & Phillips, 1985). *Diaporthe*

phaseolorum var. *caulivora* is associated with northern stem canker, and *Diaporthe phaseolorum* var. *meridionalis* is associated with southern stem canker (Fernández & Hanlin, 1996). Although both diseases cause necrosis and interveinal chlorosis, they cause other distinguishable symptoms. For northern stem canker, cankers are sunken and dark- brown, initially appear on the lower nodes, and eventually girdle the stem causing plant death. In contrast, southern stem canker lesions are restricted and rarely girdle stems (Fernandez et al., 1999).

Additional *Diaporthe* species have been postulated to cause soybean stem canker including *D. gulyea, D. eres,* and *D. longicolla*. *D. gulyea* causes reddish-brown cankers (~60 mm in length) on stems and appears to be a relatively uncommon stem canker pathogen of soybean (Mathew et al., 2018a). However, *D. gulyae* is one of two species most commonly associated with *Phomopsis* stem canker of sunflower (Mathew et al., 2015). *D. eres*, which is not a host-specific pathogen of soybean, also causes reddish-brown stem cankers on soybean (Mathew et al., 2018b). Furthermore, *D. longicolla*, the primary causal agent of PSD, can cause stem cankers that girdle soybean stems (Tolbert & Spurlock, 2017).

Species concepts of *Diaporthe* species

The species is the basic rank in the biological classification. However, fungal species can be defined differently by different mycologists, and there are diverse methods (and philosophies) associated with fungal species delineation. Attempts to universally define what constitutes a fungal species have been unsuccessful, and, consequently, many species concepts are in use (Guarro et al., 1999). However, three species concepts have been widely used to delineate *Diaporthe* species. The first species concept applied to *Diaporthe* species was the ecological species concept, which defines a species as a lineage that 1) occupies an adaptive area diverse from other lineages in its range, and 2) evolves individually from other related species (Shenoy

et al., 2007). For *Diaporthe* species, the ecological niche is most commonly another living host, which is gives rise to a host-based species concept. Applications of the host-based species concept resulted in a proliferation of named species (Gomes et al., 2013). However, observations that more than one host can be occupied by a single *Diaporthe* species, or vice versa, led to a restructuring of *Diaporthe* taxonomy based mainly the morphological species concept (Wehmeyer, 1933), in which species are defined based on overall morphological similarity among individuals (Shenoy et al., 2007). Because of high levels of morphological plasticity among *Diaporthe* spp. and their anamorphs (*Phomopsis* spp.), the morphological species concept was considered insufficient for species delimitation (Van der Aa et al., 1990; Udayanga et al., 2011). Subsequently, the phylogenetic species concept, which is a DNA sequence-based method combined with morphological characters, was applied to differentiate *Diaporthe* species (Santos et al., 2011; Udayanga et al., 2015). Within this concept, a species should be comprised of a monophyletic group of organisms that descended from a common ancestor and share at least one uniquely derived character (Moncalvo, 2005). Phylogenetic species are postulated from phylogenetic trees based on either a single gene such as the internal transcribed spacer (ITS) of nuclear ribosomal DNA (rDNA) (Pryor & Gilbertson, 2000) or multiple concatenated genes, e.g., Genealogical Concordance Phylogenetic Species Recognition (GCPSR). GCPSR provides substantially higher support to delimit *Diaporthe* species boundaries than single gene phylogenies (dos Santos et al., 2016). Therefore, species delineation via GCPSR, as supported by morphological features, is now the primary approach utilized to discriminate *Diaporthe* species.

History of *Diaporthe/Phomopsis* species associated with soybean

Diaporthe/Phomopsis species were not known to associate with soybean until the 1920s. The first report of *Diaporthe/Phomopsis* on soybean was incorrectly attributed as being a species

of *Phoma* (Wolf & Lehman, 1920), which was later named *Phomopsis sojae* by Lehman (1922). At that time, only the asexual stage was observed. However, Lehman (1923) renamed the pathogen *D. sojae* after detecting perithecia that were consistent with *Diaporthe* but differed from those of *D. phaseolorum,* the cause of pod blight on lima bean. Subsequently, Wehmeyer (1933) postulated that *D. sojae* was a variety of *D. phaseolorum,* and thus renamed it *D. phaseolorum* var. *sojae*. Athow & Caldwell (1954) postulated that a separate, distinct causal agent of soybean stem canker was a variety of *D. phaseolorum*, which they named *D. phaseolorum* var. *caulivora*. Yet another *Diaporthe*/*Phomopsis* species, *Phomopsis glycines* Petrak, was recorded by Petrak & Sydow (1936) on soybean in Japan. For decades thereafter, these were the primary *Diaporthe*/*Phomopsis* species associated with soybean worldwide.

By the late 1970s, an undescribed *Phomopsis* species was most closely associated with PSD (Kmetz et al., 1978). This fungus was later formally described as *Phomopsis longicolla* Hobbs (Hobbs et al., 1985). Although the name *D. phaseolorum* f. sp. *meridionalis* was initially suggested for the causal agent of southern stem canker (Morgan-Jones, 1989), Fernández and Hanlin (1996) formally named the fungus *D. phaseolorum* var. *meridionalis*, which was distinct from *D. phaseolorum* var. *caulivora*. Analyses of DNA sequence from the ITS region and the translation elongation factor-1 alpha (TEF1 or EF1-α) gene, in conjunction with morphological characteristics, led to the renaming of *D. phaseolorum* var. *meridionalis* as *D. aspalathi* Jansen, Castl. & Crous (van Rensburg et al., 2006). Likewise, Santos et al. (2011) renamed *P. longicolla* and *D. phaseolorum* var. *caulivora* based on multigene phylogenies as *D. longicolla* (Hobbs) J.M. Santos, Vrandečić & A.J.L. Phillips, comb. nov. and *D. caulivora* (Athow & Caldwell) J.M. Santos, Vrandečić & A.J.L. Phillips, comb. & stat. nov. respectively. Furthermore, the novel species *D. novem* J.M. Santos, Vrandečić & A.J.L. Phillips, sp. nov. was described for the first

time on soybean (Santos et al., 2011). Recently, Mathew et al. (2018b) recorded four *Diaporthe* species associated with soybean, including *D. ueckerae, D. kongii*, *Diaporthe* sp., and *D. eres*. The first three of these species caused stem diseases, whereas *D. eres* caused Phomopsis seed decay. The authors postulated that the three taxonomically identified species are not host-specific pathogens of soybean.

Fungal lifestyles

Diverse fungal lifestyles have been reported in various fungal genera, including *Diaporthe* (Gomes, et al., 2013), *Diplodia, Colletotrichum, Moniliophthora, Scleotinia,* and *Neurospora* (Kabbage et al., 2015; Rai & Agarkar, 2016). These diverse lifestyles have been broadly classified as either symbiotic, saprophytic, or a combination of the two (Rodriguez & Redman, 1997). Symbiotic lifestyles are considered by some to be more complicated than saprophytic lifestyles (Cooke & Rayner, 1984). Symbiosis can be defined as the constant relationship between two or more different organisms during at least a portion of their lifecycles (Rai & Agarkar, 2016). Many symbiotic life modes have been described, including parasitic, communalistic, and mutualistic lifestyles. In the context of plant-fungal interactions, a parasitic lifestyle is one in which a fungus gains benefits at the expense of a plant host. In a communalistic lifestyle, there is no apparent loss or gain to either organism. In a mutualistic lifestyle, both the fungus and the plant host derive one or more benefits (Rodriguez & Redman, 1997). Fungal endophytism potentially spans all of these symbiotic life modes. Oftentimes, a host plant suffers no discernable damage when colonized by an endophytic fungus, and frequently benefits from such occupation. A fine-tuned balance between the demands of the fungus and the plant's responses provides this advantage. However, if the plant-fungal

interaction becomes unbalanced, the fungus is potentially eliminated by induced host defense responses, or disease symptoms appear (Kogel et al., 2006).

Switching among symbiotic lifestyles is a result of plant-fungal communications that dictate whether interactions represent mutualism, commensalism, or parasitism (Johnson et al., 1997; Redman et al., 2001). Lifestyle alteration can be used by fungi in conjunction with other survival strategies to associate differently with various potential plant hosts (Rai & Agarkar, 2016). Consistent with this concept, *Diaporthe* species have been reported as plant pathogens, endophytes, or saprotrophs (Gomes et al., 2013). For example, *D. phaseolorum* was reported as a pathogen of soybean in Croatia (Santos et al., 2011). However, another study detected the same species as an endophyte on mangrove forests in Brazil (Sebastianes et al., 2012). Also, *D. unshiuensis* have been identified as a plant pathogen on wild asparagus (*Asparagus kiusianus*) in Japan (Dinh et al., 2019), but as an endophyte on *Carya illinoensis* in China (Yang et al., 2018). Thus, lifestyle switching is likely to be a key adaptive trait among many *Diaporthe* species.

Research rationale

Since *Diaporthe* species are associated with diseases causing substantial soybean yield losses (Allen et al., 2019), developing effective strategies to manage these destructive diseases is a priority. However, building robust strategies such as genetically resistant plants depends on the accurate identification of the underlying pathogens, understanding genetic diversity among pathogen populations, and the genetic basis of pathogenesis. Limited information about *Diaporthe* species associated with soybean is a key challenge for disease management. Therefore, this project was comprised of the following three objectives:

Objectives

1- Determine the taxonomic diversity, distribution, and pathogenic potential of *Diaporthe* species associated with asymptomatic infection of soybean in Arkansas.

2- Determine the population structure and genetic diversity of *Diaporthe* communities associated with soybean in Arkansas.

3- Identify pathogenicity genes in *D. longicolla* through a forward genetic approach.

The first objective is addressed in chapter two; the second objective is addressed in chapter three and the third objective is covered in chapter four. A fifth chapter provides concluding thoughts and potential future research directions to conclude this dissertation.

Literature Cited

Allen, T. W., Bissonnette, K., Bradley, C. A., Damicone, J. P., Dufault, N. S., Faske, T. R., . . . Young, H. (2019). *Southern United States soybean disease loss estimates for 2018.* Paper presented at the The 46th Annual Meeting of the Southern Soybean Disease Workers, March 6 – 7, Pensacola Beach, Florida.

Ash, G. J., Stodart, B., Sakuanrungsirikul, S., Anschaw, E., Crump, N., Hailstones, D., & Harper, J. D. I. (2010). Genetic characterization of a novel *Phomopsis* sp., a putative biocontrol agent for Carthamus lanatus. *Mycologia,* 102(1), 54-61.

Athow, K. L., & Caldwell, R. M. (1954). A comparative study of Diaporthe stem canker and pod and stem blight of Soybean. *Phytopathology,* 44(6), 319-325

Backman, P. A. (1985). Soybean stem canker: An emerging disease problem. *Plant Disease,* 69(8), 641. doi:10.1094/PD-69-641

Baker, D. M., Minor, H. C., Brown, M. F., & Brown, E. A. (1987). Infection of immature soybean pods and seeds by *Phomopsis longicolla*. *Canadian Journal of Microbiology,* 33(9), 797-801. doi:10.1139/m87-135

Bandara, A. Y., Weerasooriya, D. K., Bradley, C. A., Allen, T. W., & Esker, P. D. (2019). Dissecting the economic impact of soybean diseases in the United States over two decades. *bioRxiv*, 655837.

Bandre, T. R., & Šašek, V. (1977). Antibiotic activity of pyrenomycetes under submerged conditions. *Folia Microbiologica,* 22(4), 269-274.

Cooke, R. C., & Rayner, A. D. M. (1984). *Ecology of saprotrophic fungi*. New York, NY: Longman. p 415.

Dai, J., Krohn, K., Flörke, U., Gehle, D., Aust, H. J., Draeger, S., . . . Rheinheimer, J. (2005). Novel highly substituted biraryl ethers, phomosines D–G, isolated from the endophytic fungus *Phomopsis* sp. from Adenocarpus foliolosus. *European Journal of Organic Chemistry,* 2005(23), 5100-5105.

Dettrakul, S., Kittakoop, P., Isaka, M., Nopichai, S., Suyarnsestakorn, C., Tanticharoen, M., & Thebtaranonth, Y. (2003). Antimycobacterial pimarane diterpenes from the Fungus *Diaporthe* sp. *Bioorganic & Medicinal Chemistry Letters*, 13(7), 1253-1255.

Dinh, T. L., Zaw, M., & Matsumoto, M. (2019). *Diaporthe* species complex occurring on *Asparagus kiusianus* in Japan. *Journal of Plant Pathology,* 101(1), 161-167. doi:10.1007/s42161-018-0140-9

dos Santos, T. T., Leite, T. D., de Queiroz, C. B., de Araujo, E. F., Pereira, O. L., & de Queiroz, M. V. (2016). High genetic variability in endophytic fungi from the genus *Diaporthe*

isolated from common bean (*Phaseolus vulgaris* L.) in Brazil. *Journal of Applied Microbiology,* 120(2), 388-401. doi:10.1111/jam.12985

Elsäesser, B., Krohn, K., Flörke, U., Root, N., Aust, H. J., Draeger, S., . . . Kurtán, T. (2005). X-ray structure determination, absolute configuration and biological activity of phomoxanthone A. *European Journal of Organic Chemistry,* 2005(21), 4563-4570.

Fernandez, F. A., Phillips, D. V., Russin, J. S., & Rupe, J. C. (1999). Stem Canker. In G. L. Hartman, J. B. Sinclair, & J. C. Rupe (Eds.), *Compendium of Soybean Diseases* (pp. 33-35). St. Paul, MN: American Phytopathological Society.

Fernández, F. A., & Hanlin, R. T. (1996). Morphological and RAPD analyses of *Diaporthe phaseolorum* from soybean. *Mycologia*, 88(3), 425-440.

Garcia-Reyne, A., López-Medrano, F., Morales, J. M., Garcia Esteban, C., Martin, I., Erana, I., ... & Aguado, J. M. (2011). Cutaneous infection by *Phomopsis longicolla* in a renal transplant recipient from Guinea: first report of human infection by this fungus. *Transplant Infectious Disease*, 13(2), 204-207.

Gillen, A. M., Smith, J. R., Mengistu, A., & Bellaloui, N. (2012). Effects of maturity and *Phomopsis longicolla* on germination and vigor of soybean seed of near-isogenic lines. *Crop Science*, 52(6), 2757-2766. doi:10.2135/cropsci2011.10.0566

Gomes, R. R., Glienke, C., Videira, S. I., Lombard, L., Groenewald, J. Z., & Crous, P. W. (2013). *Diaporthe*: a genus of endophytic, saprobic and plant pathogenic fungi. *Persoonia,* 31, 1-41. doi:10.3767/003158513X666844

Guarro, J., Gené, J., & Stchigel, A. M. (1999). Developments in fungal taxonomy. *Clinical Microbiology Reviews,* 12(3), 454-500.

Hobbs, T. W., & Phillips, D. V. (1985). Identification of *Diaporthe* and *Phomopsis* isolates from soybean. *Phytopathology,* 75, 500.

Hobbs, T. W., Schmitthenner, A. F., & Kuter, G. A. (1985). A new *Phomopsis* species from soybean. *Mycologia,* 77(4), 535-544.

Isaka, M., Jaturapat, A., Rukseree, K., Danwisetkanjana, K., Tanticharoen, M., & Thebtaranonth, Y. (2001). Phomoxanthones A and B, novel xanthone dimers from the endophytic fungus *Phomopsis* species. *Journal of Natural Products,* 64(8), 1015-1018.

Johnson, N. C., Graham, J. H., & Smith, F. A. (1997). Functioning of mycorrhizal associations along the mutualism–parasitism continuum. *The New Phytologist,* 135(4), 575-585.

Kabbage, M., Yarden, O., & Dickman, M. B. (2015). Pathogenic attributes of *Sclerotinia sclerotiorum*: switching from a biotrophic to necrotrophic lifestyle. *Plant Sci.,* 233, 53-60. doi:10.1016/j.plantsci.2014.12.018

Kmetz, K. T. (1975). Soybean seed decay: Studies on disease cycles, effects of cultural practices on disease severity and differentiation of the pathogens *Phomopsis* sp., *Diaporthe phaseclorum* var. *sojae* and *Diaporthe phaseolorum* var. *caulivora*. Doctoral dissertation, The Ohio State University.

Kmetz, K. T., Schmitthenner, A. F., & Ellett, C. W. (1978). Soybean seed decay: Prevalence of infection and symptom expression caused by *Phomopsis* sp., *Diaporthe phaseolorum* var. *sojae*, and *D. phaseolorum* var. *caulivora*. *Phytopathology*, 68(6), 836-841.

Kobayashi, H., Meguro, S., Yoshimoto, T., & Namikoshi, M. (2003). Absolute structure, biosynthesis, and anti-microtubule activity of phomopsidin, isolated from a marine-derived fungus *Phomopsis* sp. *Tetrahedron,* 59(4), 455-459.

Kogel, K.-H., Franken, P., & Hückelhoven, R. (2006). Endophyte or parasite–what decides? *Current Opinion in Plant Biology,* 9(4), 358-363.

Kulik, M. M., & Sinclair, J. B. (1999a). Phomopsis seed decay. In G. L. Hartman, J. B. Sinclair, & J. C. Rupe (Eds.), *Compendium of soybean diseases,* (4 ed., pp.31-32). St. Paul, MN: American Phytopathological Society.

Kulik, M. M., & Sinclair, J. B. (1999b). Pod and stem blight. In G. L. Hartman, J. B. Sinclair, & J. C. Rupe (Eds.), *Compendium of soybean diseases* (4 ed., pp. 32-33). St. Paul, MN: American Phytopathological Society.

Kumaran, R. S., & Hur, B. K. (2009). Screening of species of the endophytic fungus *Phomopsis* for the production of the anticancer drug taxol. *Biotechnology and Applied Biochemistry,* 54(1), 21-30.

Lehman, S. G. (1922). Pod and stem blight of the soybean. *Journal of the Elisha Mitchell Scientific Society,* 38, 13.

Lehman, S. G. (1923). Pod and stem blight of soybean. *Annals of the Missouri Botanical Garden,* 10(2), 111-178.

Li, S. (2011). Phomopsis seed decay of soybean. In *Soybean-Molecular Aspects of Breeding*: IntechOpen.

Lin, X., Huang, Y., Fang, M., Wang, J., Zheng, Z., & Su, W. (2005). Cytotoxic and antimicrobial metabolites from marine lignicolous fungi, *Diaporthe* sp. *FEMS Microbiology Letters,* 251(1), 53-58.

Mathew, F. M., Alananbeh, K. M., Jordahl, J. G., Meyer, S. M., Castlebury, L. A., Gulya, T. J., & Markell, S. G. (2015). Phomopsis stem canker: A reemerging threat to sunflower (*Helianthus annuus*) in the United States. *Phytopathology*, 105(7), 990-997.

Mathew, F. M., Gulya, T. J., Jordahl, J. G., & Markell, S. G. (2018a). First report of stem disease of soybean (*Glycine max*) caused by *Diaporthe gulyae* in North Dakota. *Plant Disease,* 102(1), 240-240.

Mathew, F. M., Petrovic, K., Castlebury, L., Allen, T., Bergstrom, G., Bonkowski, J., . . . Johnson, M. (2018b). *Diaporthe (Phomopsis) species on soybean: Current status in the United States.* Paper presented at the The 45th Annual Meeting of the Southern Soybean Disease Workers, March 7-8, Pensacola Beach, Florida.

Mayhew, W. L., & Caviness, C. E. (1994). Seed quality and yield of early-planted, short-season soybean genotypes. *Agronomy Journal,* 86(1), 16-19.

Mengistu, A., Castlebury, L., Smith, R., Ray, J., & Bellaloui, N. (2009). Seasonal progress of *Phomopsis longicolla* infection on soybean plant parts and its relationship to seed quality. *Plant Disease,* 93(10), 1009-1018. doi:10.1094/pdis-93-10-1009

Moncalvo, J. M. (2005). Molecular systematics: Major fungal phylogenetic groups and fungal species concepts. *Evolutionary Genetics of Fungi*, 1-33.

Morgan-Jones, G. (1989). *The Diaporthe/Phomopsis complex: taxonomic considerations.* Paper presented at the World Soybean Research Conference.

Petrak, F., & Sydow, H. (1936). A small contribution to the knowledge of the mushroom flora of Japan. *Annales Mycologici*, 34, 237-51.

Pryor, B. M., & Gilbertson, R. L. (2000). Molecular phylogenetic relationships amongst *Alternaria* species and related fungi based upon analysis of nuclear ITS and mt SSU rDNA sequences. *Mycological Research*, 104(11), 1312-1321.

Rai, M., & Agarkar, G. (2016). Plant-fungal interactions: What triggers the fungi to switch among lifestyles? *Crit Rev Microbiol,* 42(3), 428-438. doi:10.3109/1040841X.2014.958052

Redman, R. S., Dunigan, D. D., & Rodriguez, R. J. (2001). Fungal symbiosis from mutualism to parasitism: who controls the outcome, host or invader? *New Phytologist,* 151(3), 705-716.

Rodriguez, R. J., & Redman, R. S. (1997). Fungal life-styles and ecosystem dynamics: biological aspects of plant pathogens, plant endophytes and saprophytes. In *Advances in Botanical Research* (Vol. 24, pp. 169-193): Elsevier.

Rossman, A. Y., Adams, G. C., Cannon, P. F., Castlebury, L. A., Crous, P. W., Gryzenhout, M., . . . Udayanga, D. (2015). Recommendations of generic names in Diaporthales competing for protection or use. *IMA Fungus,* 6(1), 145-154.

Santos, J. M., Vrandečić, K., Cosić, J., Duvnjak, T., & Phillips, A. J. (2011). Resolving the *Diaporthe* species occurring on soybean in Croatia. *Persoonia,* 27, 9-19. doi:10.3767/003158511X603719

Sebastianes, F. L., Lacava, P. T., Fávaro, L. C., Rodrigues, M. B., Araújo, W. L., Azevedo, J. L., & Pizzirani-Kleiner, A. A. (2012). Genetic transformation of *Diaporthe phaseolorum*, an endophytic fungus found in mangrove forests, mediated by *Agrobacterium tumefaciens*. *Current Genetics*, 58(1), 21-33.

Shenoy, B. D., Jeewon, R., & Hyde, K. D. (2007). Impact of DNA sequence-data on the taxonomy of anamorphic fungi. *Fungal Diversity*, 26,1-54.

Sinclair, J. B. (1993). Phomopsis seed decay of soybeans: a prototype for studying seed disease. *Plant Disease,* 77(4), 329-334.

Thompson, S. M., Tan, Y. P., Shivas, R. G., Neate, S. M., Morin, L., Bissett, A., & Aitken, E. A. B. (2015). Green and brown bridges between weeds and crops reveal novel *Diaporthe* species in Australia. *Persoonia,* 35, 39-49. doi:10.3767/003158515x687506

Ting, A. S. Y., Lee, M. V. J., Chow, Y. Y., & Cheong, S. L. (2016). Novel exploration of endophytic Diaporthe sp. for the biosorption and biodegradation of triphenylmethane dyes. *Water, Air, & Soil Pollution*, 227(4), 109.

Tolbert, A. C., & Spurlock, T. N. (2017). First Report of *Phomopsis longicolla* Causing Stem Necrosis on Soybean in Arkansas. *Plant Disease,* 101(12), 2147-2147. doi:10.1094/pdis-02-17-0223-pdn

Udayanga, D., Castlebury, L. A., Rossman, A. Y., Chukeatirote, E., & Hyde, K. D. (2015). The *Diaporthe sojae* species complex: Phylogenetic re-assessment of pathogens associated with soybean, cucurbits and other field crops. *Fungal Biology,* 119(5), 383-407. doi:10.1016/j.funbio.2014.10.009

Udayanga, D., Liu, X., McKenzie, E. H. C., Chukeatirote, E., Bahkali, A. H. A., & Hyde, K. D. (2011). The genus *Phomopsis*: biology, applications, species concepts and names of common phytopathogens. *Fungal Diversity,* 50(1), 189-225.

USDA-NASS. (2017). *Crop Production 2016 Summary*. Retrieved from https://www.nass.usda.gov

Van der Aa, H. A., Noordeloos, M. E., & Gruyter, J. D. (1990). Species concepts in some larger genera of the Coelomycetes. *Studies in Mycology* (32), 3-19.

van Rensburg, J. C. J., Lamprecht, S. C., Groenewald, J. Z., Castlebury, L. A., & Crous, P. W. (2006). Characterisation of *Phomopsis* spp. associated with die-back of rooibos (Aspalathus linearis) in South Africa. *Studies in Mycology,* 55, 65-74.

Wehmeyer, L. E. (1933). *The genus Diaporthe Nitschke and its segregates* (Vol. 9): University of Michigan Press Ann Arbor.

Williamson, N. P., Highet, A. S., Gams, W., Sivasithamparam, K., & Cowling, W. A. (1994). Diaporthe toxica sp. nov., the cause of lupinosis in sheep. *Mycological Research*, 98 (12), 1364-1368.Wolf, F. A., & Lehman, S. G. (1920). Notes on new or little known plant diseases in North Carolina in 1920. *NC Agr. Exp. Stat. Ann. Rpt,* 43, 55-58.

Wolf, F. A., & Lehman, S. G. (1920). Notes on new or little known plant diseases in North Carolina in 1920. *NC Agr. Exp. Stat. Ann. Rpt,* 43, 55-58.

Xue, A. G., Morrison, M. J., Cober, E., Anderson, T. R., Rioux, S., Ablett, G. R., . . . Zhang, J. X. (2007). Frequency of isolation of species of *Diaporthe* and *Phomopsis* from soybean plants in Ontario and benefits of seed treatments. *Canadian Journal of Plant Pathology-Revue Canadienne De Phytopathologie*, 29(4), 354-364.

Yang, Q., Fan, X. L., Guarnaccia, V., & Tian, C. M. (2018). High diversity of *Diaporthe* species associated with dieback diseases in China, with twelve new species described. *Mycokeys*(39), 97-149. doi:10.3897/*mycokeys*.39.26914

Chapter 2: Taxonomic diversity and pathogenicity of endophytic *Diaporthe* species associated with soybean in Arkansas

Abstract

Diaporthe species exist as plant pathogens, endophytes, or saprophytes, and are associated with a wide taxonomic range of plant hosts. Although some *Diaporthe* species are economically important pathogens of soybean, little is known about *Diaporthe* species that associate asymptomatically (endophytically) with soybean. To evaluate the diversity and pathogenic potential of endophytic *Diaporthe* species in Arkansas, 184 isolates were obtained from asymptomatic soybean stems from four locations within the state: Marianna, Rohwer, Stuttgart, and Keiser. Phylogenetic Bayesian inference and maximum likelihood trees constructed from a combined multilocus dataset (ITS, TEF1-α, TUB2, and CAL) identified four *Diaporthe* species associated with soybean in Arkansas, including *D. longicolla* (133 isolates), *D. unshiuensis* (41 isolates), *D. ueckerae* (8 isolates), and a putative undescribed *Diaporthe* sp. (2 isolates). Although *Diaporthe* species distribution varied throughout Arkansas, *D. longicolla* predominated at all locations. The pathogenicity of *D. unshiuensis, D. ueckerae,* and *Diaporthe sp*. was confirmed on soybean via Koch's postulates with a wounded stem assay. Additionally, a cut-seedling pathogenicity assay revealed varying levels of virulence among 114 isolates spanning the four species. This study provided the first report of *D. unshiuensis* associated with soybean worldwide. Additionally, this study revealed previously unknown levels of incidence and diversity of *Diaporthe* spp. associated with asymptomatic colonization of soybean in Arkansas.

Introduction

Diaporthe species (anamorph: *Phomopsis*) have been reported as pathogens, endophytes, and saprophytes of diverse plant hosts (Gomes et al., 2013; Santos & Phillips, 2009; Santos et al., 2011; Udayanga et al., 2015; Udayanga et al., 2014; Udayanga et al., 2011; Udayanga et al., 2012). Several *Diaporthe* species cause important diseases on soybean (*Glycine max*) in the U.S. and other major soybean production regions (Santos et al., 2011; Udayanga et al., 2014). In 2018, *Diaporthe* species associated with Phomopsis seed decay, stem canker, pod blight, and stem blight were collectively among the most damaging pathogens of soybean in Arkansas (Allen et al., 2019).

The taxonomy of soybean- associated *Diaporthe* species is not fully resolved. Historically, a host-based species concept in *Diaporthe* resulted in a proliferation of named species (reviewed by Gomes et al., 2013). However, observations that some species were associated with more than one host led Wehmeyer (1933) to restructure *Diaporthe* taxonomy based mainly on morphological features. Subsequently, the cultural and morphological characters of *Diaporthe* spp. and their anamorphs, *Phomopsis* spp., were deemed insufficient for species delimitation due to high levels of plasticity (Van der Aa et al., 1990; Wehmeyer, 1933). Consequently, DNA sequence-based methods combined with morphological characteristics have been used to differentiate *Diaporthe* species (Santos et al., 2011; Udayanga et al., 2015). Although the internal transcribed spacer (ITS) of nuclear ribosomal DNA (rDNA) has been used extensively for this purpose for many fungal genera (Rehner & Uecker, 1994), Genealogical Concordance Phylogenetic Species Recognition (GCPSR) criteria using the phylogenetic concordance of multiple unlinked genes provides substantially greater support to delimit *Diaporthe* species boundaries (dos Santos et al., 2016).

Diaporthe species complexes such as *D. sojae*, *D. nobilis,* and *D. eres* have been associated with various plant hosts (Gomes et al., 2013; Udayanga et al., 2014; Udayanga et al., 2015). Application of the GCPSR could resolve known and cryptic *Diaporthe* species complexes, especially among those where it is difficult to discriminate species boundaries by morphological features (Udayanga et al., 2012). For example, *D. longicolla* and closely related species are not yet taxonomically resolved within the *D. sojae* complex. In previous studies, *D. longicolla* was considered a member of the *P. sojae* complex and synonymous with *D. sojae* (Gomes et al., 2013). However, this suggestion needs to be confirmed by the analysis of type materials (Udayanga et al., 2015). Furthermore, *D. longicolla* is not a sister clade to *D. sojae* and its sexual stage, which could be used to discriminate these species, has not been detected. Thus, redefining the complex to which *D. longicolla* belongs using multilocus phylogeny with type materials is a significant step to obtain high resolution.

Although soybean stem diseases cause substantial losses (Allen et al., 2019), few studies have investigated *Diaporthe* species associated with soybean stems in Arkansas. The cause of stem canker in Arkansas was reported to be *D. phaseolorum* (Cke. & Ell.) Sacc. f. sp. *caulivora* (Athow & Caldwell) Kulik (*D. caulivora*) with tan cankers and red-purple margin symptoms (Hirrel & Kirkpatrick, 1986). *D. longicolla* and *D. sojae* have been associated with stem blight in Arkansas (Jackson, 2004) and, some isolates of *D. longicolla* cause stem canker (Tolbert & Spurlock, 2017). However, information about the taxonomic diversity of endophytic *Diaporthe* species associated with soybean stems is generally lacking, and the ability of endophytic species to cause symptoms on soybean has not been studied.

The aims of the present study were: 1) to identify endophytic *Diaporthe* species associated with soybean from Arkansas based on molecular phylogenetic analyses combined with phenotypic traits, and 2) determine *Diaporthe* species distribution and pathogenic potential.

Materials and Methods

Collection of *Diaporthe* isolates

Diaporthe isolates were collected from four experimental sites in Arkansas, USA, in 2015: the Rice Research and Extension Center in Stuttgart (34°28'30.25"N 91°25'7.27"W), the Lon Mann Cotton Research Station in Marianna (34°43'58.51"N 90°45'59.86"W), the Rohwer Research Station in Rohwer (33°49'26.40"N 91°16'35.20"W), and the Northeast Research and Extension Center in Kaiser (35°40'28.96"N 90° 5'13.24"W) (Figure 1). Fifty soybean plants (growth stage R1) were sampled arbitrarily from each site and five pieces (5 -10 mm) of each asymptomatic lower main stem were externally sterilized and plated on nine cm plates containing acidic potato dextrose agar (APDA, pH = 4.5). One isolate of *Diaporthe* species per plant at each site was saved, resulting in a total of 184 *Diaporthe* isolates. No *Diaporthe* isolates were recovered from 16 of the 200 plants. For single-spore purifications, each isolate was cultured on oatmeal agar (OMA) and incubated for 7-14 days at 25 ºC with a 12/12h light/dark cycle. After pycnidia formed, 50 µl of sterile, distilled water was mixed with the conidial mass produced by a single pycnidium, and the spore suspension was then transferred and spread on water agar (WA) plates. Inoculated WA plates were incubated for 24 h and a germinated single spore per each isolate was then picked and transferred to PDA plates under a microscopic field and incubated at 25 ºC in the dark for 5 days (Udayanga et al., 2012). Mycelial cultures were stored in 50% glycerol at -80 °C.

DNA extraction, PCR amplification, and sequencing

Each of the 184 *Diaporthe* isolates was grown in potato dextrose broth (PDB) at 25 °C for five days with continuous shaking (150 rpm). Mycelia were prepared for DNA extraction by rising three times with sterile water (30 ml) and centrifugation at ~ 3000 g (5000 rpm) (Sorvall Super T21, DuPont Company, Wilmington, DE, USA), followed by lyophilization for 24 h. Fungal tissue (100-200 mg) was ground with beads in a Qiagen Tissuelyser (Qiagen, CA, USA). The samples were shaken on the highest speed (30 Hz) for 2 min. Genomic DNA was extracted with a CTAB protocol (Leslie & Summerell, 2006). The quality and quantity of DNA were evaluated with a Nanodrop1000 spectrophotometer (Thermo Scientific, MA, USA). To build phylogenetic trees and delimit boundaries of *Diaporthe* species, four loci were amplified as suggested by Udayanga et al. (2012): the internal transcribed spacer (ITS) of the ribosomal DNA (rDNA); the partial translation elongation factor 1-alpha (TEF1-α) region; beta-tubulin (TUB2) region; and the partial calmodulin (CAL) region (Soares et al., 2018). PCR amplifications were performed with the ITS1 and ITS4 primers (White et al., 1990), the EF1-728F and EF1-986R primers (Carbone & Kohn, 1999), the CAL563F (or CL1F as an alternative forward primer for some isolates) and CLA2 primer sets (Udayanga et al., 2014), and the Bt2a and Bt2b primer sets (Glass & Donaldson, 1995) respectively (Table 1S). Reactions (30 µl) consisted of 1x PCR buffer, 0.4 mM dNTP, 0.4 µM of each primer, 2.5 mM $MgCl_2$, 1% DMSO, 0.4 µl of Taq polymerase and 20-30 ng template DNA. Cycling parameters for each locus were adapted from Udayanga et al. (2012). For the ITS and CAL loci, cycles consisted of an initiation step (95 °C for 5 min), 40 amplification cycles (95 °C for 30 s, 54 °C for 50 s, and 72 °C for 1 min), and a final elongation step (72 °C for 10 min). For TUB2 and TEF1-α loci, cycles consisted of an initiation step (95 °C for 5 min), 40 amplification cycles (95 °C for 30 s, 57 °C for 50 s, and 72

°C for 1 min), and a final elongation step (72 °C for 10 min). PCR products were evaluated via electrophoresis in 1% agarose gels stained with GelRed® Nucleic Acid Gel Stain (Biotium, Inc., Hayward, CA, USA). Sanger sequencing was performed by Genewiz Inc. (South Plainfield, NJ, USA). Sequence quality was assessed with Geneious software version 9.1.8 as a percentage of high quality (HQ%) (Kearse et al., 2012), and seqeunces of sufficient quality were deposited in GenBank under accession numbers: MN586627 - MN586810 for ITS, MN651137-MN651320 for TEF1-α, MN698395 - MN698578 for TUB2, and MN725831 - MN726014 for CAL (Table 2S).

Sequence alignment and phylogenetic analyses

ITS, TEF1-α, TUB2, and CAL sequences from 184 *Diaporthe* isolates were edited with Geneious version 9.1.8. Sequences of each locus were aligned with available GenBank sequences of *Diaporthe* species (with *Diaporthella corylina* as an outgroup; Table 1) using default settings of MAFFT v7.309 (Katoh & Standley, 2013), and were manually adjusted as necessary. Alignments were concatenated with Geneious version 9.1.8. A neighbor-joining tree was constructed to identify closely related taxa and reduce the number of GenBank sequences included in Bayesian analyses and maximum likelihood trees (Table 1). Bayesian analyses to infer phylogenetic trees were performed with MrBayes V3.2.6 (Huelsenbeck & Ronquist, 2001) in Geneious version 9.1.8 using the four concatenated loci with the nucleotide substitution model selected by jModeltest (Darriba et al., 2012), which is a general time- reversible (GTR) model. A Markov chain Monte Carlo (MCMC) analysis was run over 1,100,000 generations and trees were sampled every 1000 generations with the heating chain temperature set at 0.2, which resulted in 1100 trees. The first 100 trees were discarded, and the remaining 1000 trees were used to calculate posterior probabilities (PP) in the majority rule consensus tree (Andjic et al., 2016).

Maximum likelihood (ML) analyses were performed with RAxML V7.2.8 (Stamatakis, 2006) in Geneious version 9.1.8 and run with rapid bootstrapping for 1000 replicates. The RAxML software accommodated the GTR model of nucleotide substitution with the additional options of modeling gamma rate heterogeneity (G) and proportion invariable sites (I). Trees were visualized in Geneious version 9.1.8, and a tree alignment was submitted to TreeBASE (https://treebase.org/treebase-web/home.html; submission number S25789). Alpha species diversity in Arkansas locations was calculated with PC-ORD V6.19 (McCune et al., 2002).

Morphological characterization and species description

Morphological features of *Diaporthe* isolates were characterized on OMA, PDA, and autoclaved soybean stems. To investigate micromorphological structures of *Diaporthe* isolates, strains were grown on ~ 5 cm sterilized stem segments from three-week-old seedlings (autoclaved twice) of soybean cultivar 'Williams 82' on PDA for two weeks at 23 °C with a 12/12h light/dark cycle. At least 20 pycnidia, conidiophores, and conidia were measured for each isolate. The mean ($\overline{x}$), standard deviation (SD), and ranges of lengths and widths of conidia and conidiophores, as well as the diameters of pycnidia, were calculated and formatted as (min-) mean-SD – mean+ SD (- max) as described by Dinh et al. (2019). To document colony characteristics, cultures were grown on 9 cm plates of PDA and OMA for two weeks at 23 °C with a 12/12h light/dark cycle. Colony color, diameter, and appearance were recorded, as were size and shape of stromata, and arrangement of pycnidia (Thompson et al., 2011). Digital images were captured with scale bars with an Axiovert Zeiss microscope and an Axio com ICc1 camera (Carl Zeiss MicroImaging, Thornwood, NJ), a Nikon Eclipse N-U microscope (Nikon, Tokyo, Japan) with an Excelis HD camera (Unitron, Commack, NY), and a dissecting microscope connected to an Excelis HD camera (Unitron, Commack, NY).

Pathogenicity of *Diaporthe* species using the cut-stem pathogenicity assay

Pathogenicity assays were performed with 41 isolates of *D. unshiuensis,* eight isolates of *D. ueckerae*, two isolates of *Diaporthe* sp., and 63 isolates of *D. longicolla*. A cut-seedling pathogenicity assay as described by Zaccaron (2019) was utilized, with wild-type strain PL2010 of *D. longicolla* included as a positive control and mock inoculations (APDA plugs without fungal inoculation) as a negative control. Briefly, four soybean seedlings of cultivar Williams 82 were grown in pots (10 x 10 x 9 cm) for 17 days. For each pot, stems were cut and inoculated with 5 mm plugs taken from 10-day-old APDA cultures of each isolate. Plugs were inverted in barrels of pipette tips (200µL) (USA Scientific Plastics, Ocala, FL), which were affixed to cut soybean stems. After two days of incubation in a dew chamber, tips were removed and pots were transferred to a greenhouse. Pots were arranged in the greenhouse with a complete randomized block design that included three pots for each isolate. Lesion lengths were measured on stems seven days post inoculation to evaluate pathogenesis.

Koch's postulates with *Diaporthe* species on soybean

The stem-wound inoculation method is a reliable assay for *Diaporthe* pathogenesis (Ghimire et al., 2019). With this assay, pathogenicity was evaluated for three isolates of *D. unshiuensis,* three isolates of *D. ueckerae* and two isolates of *Diaporthe* sp. Wild-type strain PL2010 of *D. longicolla* was included as a positive control. The stems of three-week old soybean seedlings, c.v. Williams 82, were wounded with a sharp blade below the first trifoliate node. For inoculations, mycelial plugs (5 mm diameter) were taken from the margin of a 10-day old *Diaporthe* APDA culture, placed into fresh wounds, and secured via parafilm. The experiment utilized a complete randomized block design, with three replicates, each consisting of five plants per pot. To evaluate the ability of *Diaporthe* species to kill soybean shoots above inoculation

sites, shoot death was evaluated four days after inoculation. Ten days after inoculation, all disease symptoms were evaluated. Symptomatic stem sections (0.5 -1 cm) from fifteen plants inoculated with each isolate were externally sterilized in ethanol (70%) for 30 s and bleach (0.8%) for 1 min, rinsed 3 times with distilled sterilized water, and plated on 9 cm APDA plates (PH, 4.5). After purifying and morphologically identifying cultures, inoculations and isolations were repeated to fulfill Koch's postulates. The isolates were purified using the single spore technique described above and identified based on morphological characteristics. Morphological identification was complemented by sequencing ITS and TEF1-α loci for reisolated strains, as described above. The sequences of inoculated and recovered isolates were aligned with the default settings of MAFFT in Geneious version 9.1.8 to confirm recovery of the inoculated strain.

Results

Fungal collection

A total of 184 *Diaporthe* strains were obtained from 200 soybean plants (50 plants collected per site). Forty-eight strains (96% recovery) were collected from both Rohwer and Keiser, forty-seven strains (94% recovery) were obtained from Marianna, and forty-one strains (82% recovery) were collected from Stuttgart (Table 2).

Phylogenetic analysis

The alignment, which consisted of the combined multi-locus dataset (ITS, TEF1-α, TUB2, and CAL) from 204 strains (736 sequences of 184 strains from this study and sequences of 20 strains belonging to 17 taxa publicly available from GenBank), was comprised of 1,838 characters including alignment gaps. Of the 1,838 characters, 932 characters were parsimony-

informative, and the rest were identical sites and gaps. Isolates of *Diaporthe* species collected in this study clustered in four distinct clades, corresponding to three known species (*D. longicolla, D. unshiuensis,* and *D. ueckerae*) and one taxonomically undescribed *Diaporthe* sp. The Bayesian posterior probability ranged from 0.74 for isolates of *D. unshiuensis* to 1.00 for other species. The three known species were in sister clades with close phylogenetic relationships. The tree topologies and clades of Bayesian analysis were similar to the ML phylogenetic trees (Figure 2, 1S).

Fungal distribution and diversity

Of the 184 *Diaporthe* strains obtained in this study, 133 isolates were *D. longicolla*, 41 isolates were related to *D. unshiuensis*, 8 isolates were related to *D. ueckerae*, and 2 isolates were related to *Diaporthe* sp. (Table 2). *D. longicolla* was predominant at all sites, with percent incidence ranging from 52% to 100%. *D. unshiuensis* represented 40.4, 8.3, and 37.5% of the isolates from Marianna, Keiser, and Rohwer, respectively; it was not collected from Stuttgart. *D. ueckerae* represented 8.3% of the isolates collected from both Keiser and Rohwer. One isolate of *Diaporthe* sp. was collected from Marianna and Rohwer. Rohwer was the most diverse location based on diversity indices with all four species. In contrast, Stuttgart was the least diverse location with the single species, *D. longicolla.*

Isolate characterization

D. unshiuensis F. Huang, Hyde K. D. & H.Y. Li sp. nov. (Figure 3, Table 3S).

MycoBank: 810845.

Etymology: named according to the species epithet of the host, *Citrus unshiu*.

Holotype: China, Zhejiang Province, Linhai, on melanose fruit of *C. unshiu*, 2009, G.Q. Chen and F. Huang, ZJUD52H (holotype, dried culture), ex-type living culture ZJUD52= CGMCC 3.17569.

Distribution: China (Huang et al., 2015; Yang et al., 2018), Japan (Dinh et al., 2019), and the USA (this study).

Sexual morphology: undetermined.

Specimens examined: Pure cultures were isolated from symptomless stems of soybean.

Culture characters: On PDA, specimens covered 9 cm petri dishes in 5 days. Colonies were circular form, raised elevation, white at first and turning to grey with aging. The surfaces of colonies were flattened, dense, and velvety with white aerial mycelia. Their reverses were off-white to gray at the center, with scattered dark brown patches of mycelium developing with age into stroma or conidiomata and/or centered patches in two-week-old cultures. On OMA, specimens covered 9 cm petri dishes in 5 days. Colonies were initially white and became zonate with 2-4 zones, with a circular form, raised elevation, and with filiform margin. Cultures had a fluffy mycelium with a few small. dark brown to black stroma near colony centers or edges that developed into pycnidial conidiomata.

Asexual micro-morphological characters: On autoclaved soybean stems, conidiomata were embedded in stem tissues and masses of scattered pycnidia on the substrate were solitary or clustered into groups of 2-3, globose to sub-globose, erumpent, white to gray, (-306) 356.6 - 517.5 (-637) μm diam (n= 30). Pycnidial walls composed of several layers of parenchymatous tissue, brownish to dark brown with texture angularis. Pycnidia had single, circular ostioles with long dark brown to black necks, 128-255 μm in size. Conidiophores were simple, hyaline, smooth, cylindrical, elongate, tightly packed, tapered to the apex, unbranched, terminal, (-7) 7.3-

13 (-16) x (-1) 1.28-2 (-3) µm (*n*= 30), and terminated with hyaline, cylindrical conidiogenous cells that formed masses of alpha conidia. Alpha conidia were hyaline, smooth, ellipsoidal, clavate and cylindrical in shape, aseptate, slightly tapered to the apex with the base truncate and truncate in both the apex and base of the small alpha conidia, often biguttulate, few triguttulate, (-5) 5.9-7.2 (-8) x (-2) 1.9-2.1 (-2.5) µm. Beta conidia were seen rarely.

Diaporthe ueckerae Udayanga & Castl. sp. nov. (Figure 4, Table 4S).

MycoBank: MB 810794.

Etymology: Named honor of the mycologist Francis A. Uecker, who was greatly involved in the collections of cultures and specimens, taxonomy, and phylogeny of the genus *Diaporthe.*

Type specimens: USA, Oklahoma, on the crown of *Cucumis melo*, F.A. Uecker (BPI 748011–holotype, dried culture), ex-type living culture FAU656 = CBS 139283.

Distribution: China, USA

Specimens examined: Pure cultures were isolated from symptomless stems of soybean.

Culture characters: On PDA, specimens covered 9 cm petri dishes within 5 days. Colonies were circular, with raised elevation and white aerial mycelium, and off-white, fluffy colony surfaces and lighter brown reverse with darkly pigmented mycelium and very few pycnidia in some isolates. Mycelia developed zones at the bottom of plates with 2-4 zones. On OMA, specimens covered 9 cm petri dishes within 5 days. Colonies were circular and dense, with raised elevation and filiform margins, fluffy mycelium, zonate growth with 2-3 zones, ranging from many small dark-brown to black stroma to scattered pycnidia and few large, black pycnidia near colony centers.

Sexual morphology: undetermined.

Asexual micro-morphological characters: Conidiomata embedded in soybean stem tissues were pycnidial and erumpent at maturity, (-128) 171-249 (-280) µm diam (*n*= 30), with solitary distribution or clustered in groups of 2-5, globose-sub globose, yellowish to gray. Conidiomata walls were parenchymatous with several layers and light to dark brown texture angularis, with elongated dark brown to black necks, 80-255 µm in length. Conidiophores were hyaline, smooth, cylindrical, slender, unbranched, (-5) 8.5-14.8 (-16) x (-1) 1.4-2 (-2) µm (*n*= 30), terminated with hyaline, cylindrical, phialidic conidiogenous cells, forming abundant alpha conidia. Alpha conidia were hyaline, smooth, aseptate, ellipsoidal shape, often biguttulate, truncate in apex and base subtruncate, (-6) 6.2-7.39 (-8.5) x (-2) 1.97-2.375 (-2.5) µm (*n*= 30). Beta conidia were seen rarely.

Diaporthe longicolla (Hobbs) J.M. Santos, Vrandečić & A.J.L. Phillips, comb. nov. (Figure 5, Table 5S).

MycoBank MB563213; Basionym. *Phomopsis longicolla* Hobbs, Mycologia 77: 542. 1985.

Distribution: Australia, Croatia, Greece, New Mexico, China, USA (Arkansas, Iowa, Illinois, Missouri, Mississippi, Nebraska, Ohio).

Specimens examined: Pure cultures were isolated from symptomless stems of soybean.

Culture characters: On PDA, specimens covered 9 cm petri dishes in 5 days. Colonies were circular, somewhat flat with filamentous margins, floccose, with densely abundant white hyphae. Colony surfaces were initially white, developing greenish to yellowish zonate growth with 2-3 zones with\dark olive to black patches of mycelium or stroma developing with age. Pycnidia were present on the surface and bottom of cultures, or very small stroma extended across the entire surface and bottom of cultures. Colony reverse was initially off-white to light pink, becoming dark olive to black with age, with massive effuse black zonate conidiomata. On OMA,

specimens covered 9 cm petri dishes in 5 days. Colonies were circular with raised elevation and few fluffy mycelia, zonate with 2-4 zones. Mycelia were initially white, developing a light olive color across the entire colony surface and a dark olive color at colony centers with yellow spots in some isolates. The central and marginal small dark olive to black patches of mycelia or stroma later developed into conidiomata.

Sexual morphology: undetermined.

Asexual micro-morphological characters: On autoclaved soybean stems, conidiomata were pycnidial, embedded in stem tissues and erumpent at maturity, solitary or clustered in groups of 2-5, globose-subglobose, gray to black and sometimes creamy, (-255) 312.9-474.3 (-510) μm, with long simple and sometimes branched dark brown to black necks,102-350 μm long, with apical ostioles. Conidiomata were composed of parenchymatous walls containing several layers of light to dark brown tissues, texture angularis. Conidiophores were hyaline, smooth, elongate, cylindrical to subcylindrical, simple with few branches, (-7) 8.9-14.3 (-18) x (-1) 1.3-1.9 (-2) μm (*n*= 30), septate with 1-2 cells, attached at apexes with hyaline conidiogenous cells. Conidiogenous cells were phialidic and cylindrical, slightly tapering towards apexes, forming hyaline alpha conidia. Alpha conidia, abundant in culture media and in soybean stem cultures, were smooth, aseptate, and ellipsoidal, often biguttulate, but the clavate shape with 3-5 guttulate cells could also be seen, acute in apex and base, sub-truncate or obtuse, (-5) 5.53 -6.4 (-7) x (-2) 1.93- 2.12 (-2.5) μm(*n*= 30). Beta conidia were seen rarely.

Virulence and pathogenicity of *Diaporthe* species

Although the *Diaporthe* isolates in this study were recovered from symptomless soybean stems, all 114 isolates evaluated, representing four distinct species, caused lesions on soybean stems in cut-stem pathogenicity assays (Figure 7). Lesion lengths were significantly different

among all isolates evaluated. However, when assessed at the species level, lesion lengths overlapped across the three taxonomically defined *Diaporthe* species (*D. longicolla, D. unshiuensis,* and *D. ueckerae*).

Fulfilling Koch's postulates of pathogenicity of *Diaporthe* species on soybean

By fulfilling Koch's postulates, *D. unshiuensis, D. ueckerae*, and *Diaporthe* sp. were confirmed to be soybean pathogens. Pure cultures were isolated from symptomatic stems of soybean after two rounds of inoculation and isolation (Figure 2S). Isolates caused stem cankers, with symptoms appearing 2-10 days post-inoculation. Symptoms included brown to reddish-brown lesions, changing to dark brown to black over time. Lesions were initially oval to ellipsoid, later extending to elongated cracks over the wounds and penetrating deeply as brown sunken cankers. Infected stem tissue decayed rapidly (Figure 8) ultimately leading to death of infected plants (Figure 9). Isolate recovery was confirmed with alignments of ITS and TEF1-α sequences from inoculated and recovered isolates (Figure 3S, 4S).

Discussion

Although accurate pathogen identification is a crucial step to breed resistant plants (Agrios, 2005), control plant diseases, and perform international phytosanitary processes (Santos et al., 2017), it is difficult to accurately identify species of *Diaporthe* with various parameters, such as morphological features (Udayanga et al., 2015), host association (Rehner & Uecker, 1994), and molecular techniques based on rDNA exclusively (Udayanga et al., 2012). Therefore, results herein agreed with previous studies that phylogenetic trees utilizing four loci successfully discriminated species of *Diaporthe* (Dissanayake et al., 2017a; 2017b). Species complexes of *Diaporthe* have been reported to be associated with field crops such as soybean (*Glycine max*)

(Udayanga et al., 2015), sunflower (*Helianthus annuus*) (Thompson et al., 2011), and Japanese wild asparagus (*Asparagus kiusianus*) (Dinh et al., 2019). Most of these fungi are important plant pathogens. In fact, genealogical concordance phylogenetic species recognition (GCPSR) was used to reconstruct phylogenetic trees and accurately discriminate cryptic species in important plant pathogenic genera (Shivas & Cai, 2012; Udayanga et al., 2014).

The phylogenetic tree presented herein could also reflect divergent events and speciation among the species studied. The timeline of speciation among the three closely related *Diaporthe* species may have begun with *D. ueckerae* diverging first, followed by *D. unshiuensis*, and lastly with *D. longicolla.*

The three taxonomically identified endophytic *Diaporthe* species in Arkansas collectively had a relatively high occurrence (82-96%) on healthy soybean stems, comparatively higher than their occurrence on soybean stems in Canada (73%) (Xue et al., 2007). The endophytic association of *Diaporthe* species could be valuable for fungal survival and epidemiological transitions; propagules of endophytes can be inoculation sources for parasitic lifestyles (Rai & Agarkar, 2016). Therefore, endophytic *Diaporthe* species of soybean could be an inoculation source for late-stage soybean diseases such as PSD, pod and stem blight, and stem canker. Additionally, findings herein also agreed with a previous study that the distribution of *Diaporthe* species sites varied among different study sites (Xue et al., 2007).

Diaporthe longicolla, one of three described *Diaporthe* species in the present study, was the most dominant species associated with soybean in a previous study (Xue et al., 2007). *D. longicolla* is known to be associated with various soybean tissues (Mengistu et al., 2009), has been reported as the primary cause of *Phomopsis* seed decay (Li, 2011; Shan et al., 2012), and has been associated with pod and stem blight (Cui et al., 2009), stem canker (Tolbert &

Spurlock, 2017), and leaf spot of soybean (Xue et al., 2015). In Arkansas, *D. longicolla* was recovered from necrotic soybean stems in 2017 (Tolbert & Spurlock, 2017) and from infected soybean seeds in 2018, with incidences ranging from 8 to 76% (Rupe et al., 2019). *D. longicolla* is also commonly isolated as an endophyte or a pathogen from a wide range of plant hosts (Li, 2011; Mengistu et al., 2007). The ubiquitous distribution of *D. longicolla* could indicate that this fungus evolved genetic mechanisms during its speciation that support diverse lifestyles.

This is the first report of *D. unshiuensis* associated with soybean worldwide, although it has been reported on other plants in China and Japan (Dinh et al., 2019; Huang et al., 2015; Yang et al., 2018). This species was described first on fruits of citrus (*Citrus unshiu*) with unidentified symptoms and as an asymptomatic endophyte on branches and twigs of kumquat (*Fortunella margarita*) (Huang et al., 2015). More recently, it has been reported from asymptomatic twigs of *Carya illinoensis* in China (Yang et al., 2018) and symptomatic stems of Japanese wild asparagus (*Asparagus kiusianus*) (Dinh et al., 2019). The significance of this species in the present study is that it occurred on soybean with a high percentage, second only to the closest phylogenetically related species *D. longicolla.* Furthermore, the pathogenicity of *D. unshiuensis* was confirmed and the virulence of the isolates evaluated was not significantly different from the virulence of *D. longicolla* isolates. Taxonomically, Huang et al. (2015) confirmed that *D. unshiuensis* is a new species distinguished from *D. longicolla* by Length/Width (L/W) ratios of alpha conidia. The mean alpha-conidium ratio of *D. unshiuensis* was smaller (2.4) than the alpha-conidium ratio of *D. longicolla* (2.9-3.4). However, in another study, the L/W ratios of alpha conidia of *D. unshiuensis* ranged from 3.1-3.4, which suggested overlap of these parameters between the two species (Yang et al., 2018). The results of this study were consistent with results presented by Yang et al. (2018) (Figure 6). In contrast to the type materials described

by Huang et al., (2015), observations in the present study discriminated isolates of *D. unshiuensis* by development of raised fluffy aerial growth of mycelium on the surface of OMA plates and lack of massive emerged stromata in the bottoms of PDA cultures. Consequently, our results confirmed that *D. unshiuensis* and *D. longicolla* are phylogenetically and morphologically distinct species although they share some overlap in morphological features.

Even though *D. unshiuensis* was not identified in the U.S. previously, ITS sequences of isolates collected in this study were identical with sequences of *Diaporthe* isolates previously isolated from soybean in the U.S. For example, the ITS sequence of unidentified *Phomopsis* isolate STAM 73 (GenBank. accession no. FJ785440) recovered from a soybean stem (V6) in Mississippi (Mengistu et al., 2009) has 100% sequence identity to *D. unshiuensis* isolates from the current study and 99.8% sequence identity to the ex-type of *D. unshiuensis* (GenBank. accession no. KJ490587.1). Moreover, ITS sequences of *D. unshiuensis* isolates in the present study are identical (100%) to the sequence of isolate DL5 (GenBank. accession no. MF125057.1), which was isolated from stem canker symptoms of soybean in Arkansas and identified as *D. longicolla* (Tolbert & Spurlock, 2017). This may indicate previous misidentifications of *D. unshiuensis*, thus underscoring the difficulty to identify *Diaporthe* species. However, the isolates mentioned above have not been thoroughly evaluated because sequences of other loci such as TEF1-α, TUB2, and CAL are not currently available.

D. ueckerae was isolated as an asymptomatic endophyte from two sites in Arkansas with a low percentage of occurrence. *D. ueckerae* has not been previously reported in association with soybean in Arkansas, although other studies reported this species as a pathogen of soybean in other U.S. production areas (Mathew et al., 2018) and in Brazil (Gomes et al., 2013; Udayanga et al., 2014). *D. ueckerae* has also been reported as pathogenic on other plant hosts, including

peanuts (*Arachis hypogaea* L. cvs. Holt, Kairi) (Thompson et al., 2018) and mango (*Mangifera indica* L.) (Lim et al., 2019). *D. ueckerae* could potentially be an opportunistic pathogen of soybean, which might explain the low incidence of this fungus in the current study.

The observation that *Diaporthe* species isolated from asymptomatic soybean stems could cause lesions in pathogenicity assays suggested that these *Diaporthe* species may have different lifestyles on soybean and are capable of switching between these lifestyles. A previous study indicated that many endophytic fungi caused disease symptoms on *Arabidopsis thaliana* in conditions stressful for the host and favorable for the endophytes, presumably due to disturbance of the delicate balance between fungal virulence and host defense that seems to be necessary for asymptomatic colonization (Junker et al., 2012). For *Diaporthe* species, Roy et al. (1997) demonstrated that isolates of *P. longicolla* (*D. longicolla*) recovered from asymptomatic leaves and stems of weedy plant species caused lesions and formed pycnidia on inoculated soybean stems. Additionally, *D. unshiuensis* has been described as an endophyte of *Fortunella margarita*, *Citrus unshiu* (Huang et al., 2015), and *Carya illinoensis* (Yang et al., 2018), but was later shown to be pathogenic on *Asparagus kiusianus* (Dinh et al., 2019). However, the underlying mechanisms of lifestyle transitions and triggers of such transitions are still unknown.

The symptomless endophytes in this study may utilize one or more of the following lifestyle strategies: commensalism, mutualism, or parasitism (Delaye et al., 2013). Therefore, more than one scenario could potentially explain the dynamics of their interactions with soybean, including shifting between asymptomatic endophytism and pathogenesis. However, such inferences need to be investigated and confirmed. The first inference for this process could be a hemibiotrophic lifestyle. This lifestyle includes features of both biotrophs and necrotrophs, in which a fungus initially invades living host cells and subsequently shifts to a necrotrophic

lifestyle to obtain nutrients from host cells that they kill (Kabbage et al., 2015). The other two scenarios are that these species may behave as commensals or mutualists and then switch at the appropriate time to a necrotrophic lifestyle. All these scenarios should be researched and dissected experimentally.

In conclusion, this study not only added *D. unshiuensis* as an unknown pathogen on soybean worldwide and as an unreported fungus in the U.S., but it also highlighted that this species along with *D. ueckerae* could share the economic impact with the prevalent pathogen, *D. longicolla,* due to their close morphological and phylogenetical relationships. Furthermore, these pathogenic species appear to colonize soybean tissues as one of their life modes without causing discernable effects to the plant host.

Literature Cited

Agrios, G. N. (2005). *Plant pathology*: Edition 5th Academic Press, New York, United States of America. 922 p.

Allen, T. W., Bissonnette, K., Bradley, C. A., Damicone, J. P., Dufault, N. S., Faske, T. R., . . . Young, H. (2019). *Southern United States soybean disease loss estimates for 2018.* Paper presented at the The 46 th Annual Meeting of the Southern Soybean Disease Workers, March 6 –7,Pensacola Beach, Florida.

Andjic, V., Maxwell, A., Hardy, G. E. S., & Burgess, T. I. (2016). New cryptic species of *Teratosphaeria* on *Eucalyptus* in Australia. *IMA Fungus,* 7(2), 253-263. doi:10.5598/imafungus.2016.07.02.05

Carbone, I., & Kohn, L. M. (1999). A method for designing primer sets for speciation studies in filamentous ascomycetes. *Mycologia,* 91(3), 553-556.

Cui, Y. L., Duan, C. X., Wang, X. M., Li, H. J., & Zhu, Z. D. (2009). First report of *Phomopsis longicolla* causing soybean stem blight in China. *Plant Pathology,* 58(4), 799-799. doi:10.1111/j.1365-3059.2009.02057.x

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods,* 9(8), 772.

Delaye, L., Garcia-Guzman, G., & Heil, M. (2013). Endophytes versus biotrophic and necrotrophic pathogens-are fungal lifestyles evolutionarily stable traits? *Fungal Diversity,* 60(1), 125-135. doi:10.1007/s13225-013-0240-y

Dinh, T. L., Zaw, M., & Matsumoto, M. (2019). *Diaporthe* species complex occurring on *Asparagus kiusianus* in Japan. *Journal of Plant Pathology,* 101(1), 161-167. doi:10.1007/s42161-018-0140-9

Dissanayake, A. J., Camporesi, E., Hyde, K. D., Wei, Z., Yan, J. Y., & Li, X. H. (2017a). Molecular phylogenetic analysis reveals seven new *Diaporthe* species from Italy. *Mycosphere*, 8(5), 853-877.

Dissanayake, A. J., Phillips, A. J. L., Hyde, K. D., Yan, J. Y., & Li, X. H. (2017b). The current status of species in *Diaporthe*. *Mycosphere*, 8(5), 1106-1156.

dos Santos, T. T., Leite, T. D., de Queiroz, C. B., de Araujo, E. F., Pereira, O. L., & de Queiroz, M. V. (2016). High genetic variability in endophytic fungi from the genus *Diaporthe* isolated from common bean (*Phaseolus vulgaris* L.) in Brazil. *Journal of Applied Microbiology,* 120(2), 388-401. doi:10.1111/jam.12985

Ghimire, K., Petrovic, K., Kontz, B. J., Bradley, C. A., Chilvers, M. I., Mueller, D. S., . . . Mathew, F. M. (2019). Inoculation Method Impacts Symptom Development Associated

with *Diaporthe aspalathi*, *D. caulivora*, and *D. longicolla* on Soybean (Glycine max). *Plant Disease,* 103(4), 677-684. doi:10.1094/pdis-06-18-1078-re

Glass, N. L., & Donaldson, G. C. (1995). Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. *Appl. Environ. Microbiol*., 61(4), 1323-1330.

Gomes, R. R., Glienke, C., Videira, S. I., Lombard, L., Groenewald, J. Z., & Crous, P. W. (2013). *Diaporthe*: a genus of endophytic, saprobic and plant pathogenic fungi. *Persoonia,* 31, 1-41. doi:10.3767/003158513X666844

Hirrel, M. C., & Kirkpatrick, T. L. (1986). First report of stem canker on soybeans in Arkansas. *Plant Disease,* 70, 78.

Huang, F., Udayanga, D., Wang, X., Hou, X., Mei, X., Fu, Y., . . . Li, H. (2015). Endophytic *Diaporthe* associated with Citrus: A phylogenetic reassessment with seven new species from China. *Fungal Biol,* 119(5), 331-347. doi:10.1016/j.funbio.2015.02.006

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics,* 17(8), 754-755.

Jackson, E. W. (2004). *Resistance to Phomopsis seed decay and purple seed stain in soybean, and virulence differences among Phomopsis spp. causing seed decay.* Doctoral dissertation. University of Arkansas, Fayetteville, Retrieved from http://uark.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwdV3fS8MwED7Uvgx8UHT4a3J_gBtdl6zL69zWKQxEfB9Jk6iobVm3wd78072k6xRRyFMSjuToJfmu390B9KJO2P51JnQZ17EVTDCmw8goKwg8a6kGwghufU7XZNYbj-L7aX_0nTNp92dXLt52rBrnuYgY4fAB34fAJUuJDiAY3iXDxDGZpa-wUt8QkyNozGS2qmzwGPZMdgKfj6Z0rzNSKy5zfHjJP_KCerCkGwO1SeUGCcdjQXt9N1Wvj2ZCamW-UUZmN37G nWx8sFBWNczIetGXyfop9ii6GAqHZP92Us7hdZk_HQ7bbvVzreOmnm9q27UhEPpCO7Z0gfC6TNA3jeELCyPhVIsVVISUAqtIE2GhpEZnkPzb2EX_w1cQqOioTh_whUElr5306o0eL1V6BdpvIae (Book, Whole)

Junker, C., Draeger, S., & Schulz, B. (2012). A fine line–endophytes or pathogens in Arabidopsis thaliana. *Fungal Ecology,* 5(6), 657-662.

Kabbage, M., Yarden, O., & Dickman, M. B. (2015). Pathogenic attributes of *Sclerotinia sclerotiorum*: switching from a biotrophic to necrotrophic lifestyle. *Plant Sci.,* 233, 53-60. doi:10.1016/j.plantsci.2014.12.018

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution,* 30(4), 772-780.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., . . . Duran, C. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics,* 28(12), 1647-1649.

Leslie, J. F., & Summerell, B. A. (2006). *The fusarium laboratory manual*. Ames, Iowa: Blackwell Pub.

Li, S. (2011). Phomopsis seed decay of soybean. In *Soybean-Molecular Aspects of Breeding*: (Sudaric A, ed.). IntechOpen 277–92. Rijeka, Croatia.

Li, S. X., Hartman, G. L., & Boykin, D. L. (2010). Aggressiveness of *Phomopsis longicolla* and Other *Phomopsis* spp. on Soybean. *Plant Disease,* 94(8), 1035-1040. doi:10.1094/pdis-94-8-1035

Lim, L., Mohd, M. H., & Zakaria, L. (2019). Identification and pathogenicity of *Diaporthe* species associated with stem-end rot of mango (Mangifera indica L.). *European Journal of Plant Pathology*, 155(2), 687-696.

Mathew, F., Petrovic, K., Castlebury, L., Allen, T., Bergstrom, G., Bonkowski, J., . . . Johnson, M. (2018). *Diaporthe (Phomopsis) species on soybean: current status in the United States.* Paper presented at the The 45thAnnual Meeting of the Southern Soybean Disease Workers, March 7-8, Pensacola Beach, Florida.

McCune, B., Grace, J. B., & Urban, D. L. (2002). *Analysis of ecological communities* MjM software design. Vol. 28. Gleneden Beach, OR.

Mengistu, A., Castlebury, L., Smith, R., Ray, J., & Bellaloui, N. (2009). Seasonal Progress of *Phomopsis longicolla* Infection on Soybean Plant Parts and Its Relationship to Seed Quality. *Plant Disease,* 93(10), 1009-1018. doi:10.1094/pdis-93-10-1009

Mengistu, A., Castlebury, L. A., Smith, J. R., Rossman, A. Y., & Reddy, K. N. (2007). Isolates of *Diaporthe-Phomopsis* from weeds and their effect on soybean. *Canadian Journal of Plant Pathology-Revue Canadienne De Phytopathologie,* 29(3), 283-289.

Rai, M., & Agarkar, G. (2016). Plant-fungal interactions: What triggers the fungi to switch among lifestyles? *Crit Rev Microbiol,* 42(3), 428-438. doi:10.3109/1040841X.2014.958052

Rehner, S. A., & Uecker, F. A. (1994). Nuclear ribosomal internal transcribed spacer phylogeny and host diversity in the coelomycete *Phomopsis*. *Canadian Journal of Botany*, 72(11), 1666-1674.

Roy, K. W., Ratnayake, S., & McLean, K. (1997). Colonization of weeds by *Phomopsis longicolla*. *Canadian Journal of Plant Pathology,* 19(2), 193-196.

Rupe, J. C., Rojas, J. A., Holland, R., Segalin, S. R., Bond, R. D., & Still, J. A. (2019). *Effect of delayed harvest on the seed quality of commercial soybean cultivars.* Paper presented at the the 46thAnnual Meeting of the Southern Soybean Disease Workers, March 6 – 7,Pensacola Beach, Florida.

Santos, J. M., & Phillips, A. J. L. (2009). Resolving the complex of *Diaporthe* (*Phomopsis*) species occurring on *Foeniculum vulgare* in Portugal. *Fungal Diversity,* 34(11), 111-125.

Santos, J. M., Vrandečić, K., Cosić, J., Duvnjak, T., & Phillips, A. J. (2011). Resolving the *Diaporthe* species occurring on soybean in Croatia. *Persoonia,* 27, 9-19. doi:10.3767/003158511X603719

Santos, L., Alves, A., & Alves, R. (2017). Evaluating multi-locus phylogenies for species boundaries determination in the genus *Diaporthe*. *Peer J.,* 5, e3120.

Shan, Z., Li, S., Liu, Y., Yang, Z., Yang, C., Sha, A., . . . Zhou, X. A. (2012). First Report of Phomopsis Seed Decay of Soybean Caused by *Phomopsis longicolla* in South China. *Plant Disease,* 96(11), 1693-1693. doi:10.1094/pdis-04-12-0401-pdn

Shivas, R. G., & Cai, L. (2012). Cryptic fungal species unmasked. *Microbiology Australia* (1) 36-37.

Soares, D. A., de Oliveira, D. P., Dos Santos, T. T., Marson, P. G., & Pimenta, R. S. (2018). Multiloci identification of *Diaporthe* fungi isolated from the medicinal plant Costus spiralis (Jacq.) Roscoe (Costaceae). *J. App.l Microbiol.,* 125(1), 172-180. doi:10.1111/jam.13769

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics,* 22(21), 2688-2690.

Thompson, S. M., Grams, R. A., Neate, S. M., Shivas, R. G., Ryley, M. J., Tan, Y. P., . . . O'Connor, D. J. (2018). First Reports of *Diaporthe kongii*, *D. masirevicii*, and *D. ueckerae* Associated with Stem and Peg Dieback on Peanut in Australia. *Plant Disease,* 102(7), 1459-1459. doi:10.1094/pdis-12-17-1930-pdn

Thompson, S. M., Tan, Y. P., Young, A. J., Neate, S. M., Aitken, E. A. B., & Shivas, R. G. (2011). Stem cankers on sunflower (*Helianthus annuus*) in Australia reveal a complex of pathogenic *Diaporthe* (*Phomopsis*) species. *Persoonia: Molecular Phylogeny and Evolution of Fungi,* 27, 80.

Tolbert, A. C., & Spurlock, T. N. (2017). First Report of *Phomopsis longicolla* Causing Stem Necrosis on Soybean in Arkansas. *Plant Disease,* 101(12), 2147-2147. doi:10.1094/pdis-02-17-0223-pdn

Udayanga, D., Castlebury, L. A., Rossman, A. Y., Chukeatirote, E., & Hyde, K. D. (2015). The *Diaporthe sojae* species complex: Phylogenetic re-assessment of pathogens associated

with soybean, cucurbits and other field crops. *Fungal Biology,* 119(5), 383-407. doi:10.1016/j.funbio.2014.10.009

Udayanga, D., Castlebury, L. A., Rossman, A. Y., & Hyde, K. D. (2014). Species limits in *Diaporthe*: molecular re-assessment of *D. citri*, *D. cytosporella*, *D. foeniculina* and *D. rudis*. *Persoonia,* 32, 83-101. doi:10.3767/003158514X679984

Udayanga, D., Liu, X., McKenzie, E. H. C., Chukeatirote, E., Bahkali, A. H. A., & Hyde, K. D. (2011). The genus *Phomopsis*: biology, applications, species concepts and names of common phytopathogens. *Fungal Diversity,* 50(1), 189-225.

Udayanga, D., Liu, X. Z., Crous, P. W., McKenzie, E. H. C., Chukeatirote, E., & Hyde, K. D. (2012). A multi-locus phylogenetic evaluation of *Diaporthe* (*Phomopsis*). *Fungal Diversity,* 56(1), 157-171. doi:10.1007/s13225-012-0190-9

Van der Aa, H. A., Noordeloos, M. E., & Gruyter, J. d. (1990). Species concepts in some larger genera of the Coelomycetes. *Studies in Mycology* (32), 3-19.

Wehmeyer, L. E. (1933). *The genus Diaporthe Nitschke and its segregates* (Vol. 9): University of Michigan Press: Ann Arbor.

White, T. J., Bruns, T., Lee, S., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: A Guide to Methods and Applications,* 18(1), 315-322.

Xue, A. G., Morrison, M. J., Cober, E., Anderson, T. R., Rioux, S., Ablett, G. R., . . . Zhang, J. X. (2007). Frequency of isolation of species of *Diaporthe* and *Phomopsis* from soybean plants in Ontario and benefits of seed treatments. *Canadian Journal of Plant Pathology-Revue Canadienne De Phytopathologie,* 29(4), 354-364.

Xue, C. S., Lu, Y. Y., Xiao, S. Q., & Duan, Y. X. (2015). First Report of *Phomopsis longicolla* Causing Leaf Spot on Soybean in China. *Plant Disease,* 99(2), 290-290.

Yang, Q., Fan, X. L., Guarnaccia, V., & Tian, C. M. (2018). High diversity of *Diaporthe* species associated with dieback diseases in China, with twelve new species described. *Mycokeys* (39), 97-149. doi:10.3897/mycokeys.39.26914

Figures and Tables

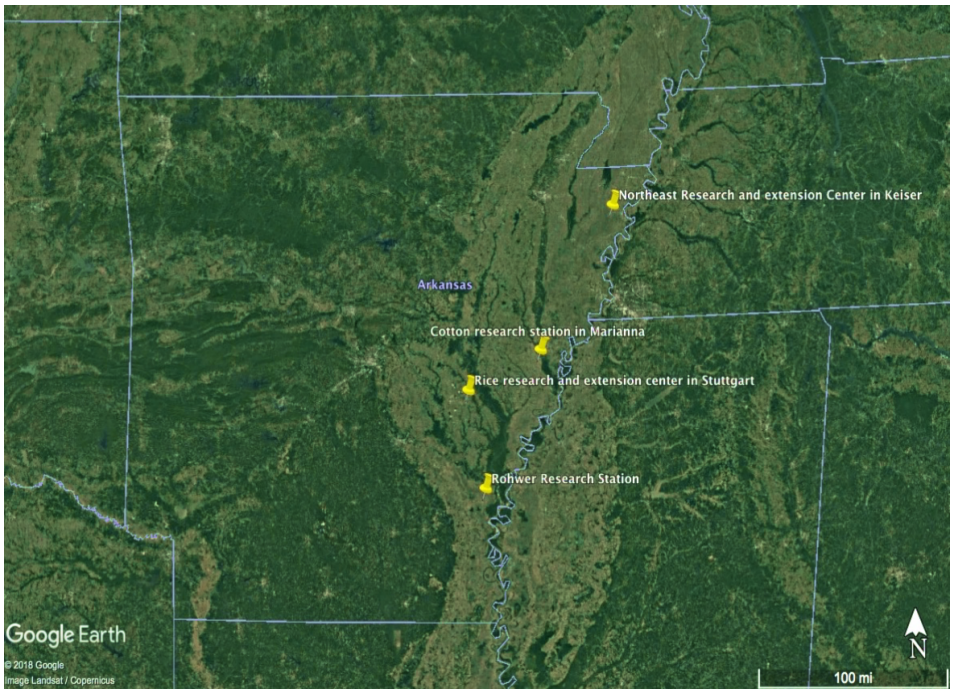


Figure 1. Locations of sampling sites in Arkansas, U.S.A. (Google Earth © 2018 Google).

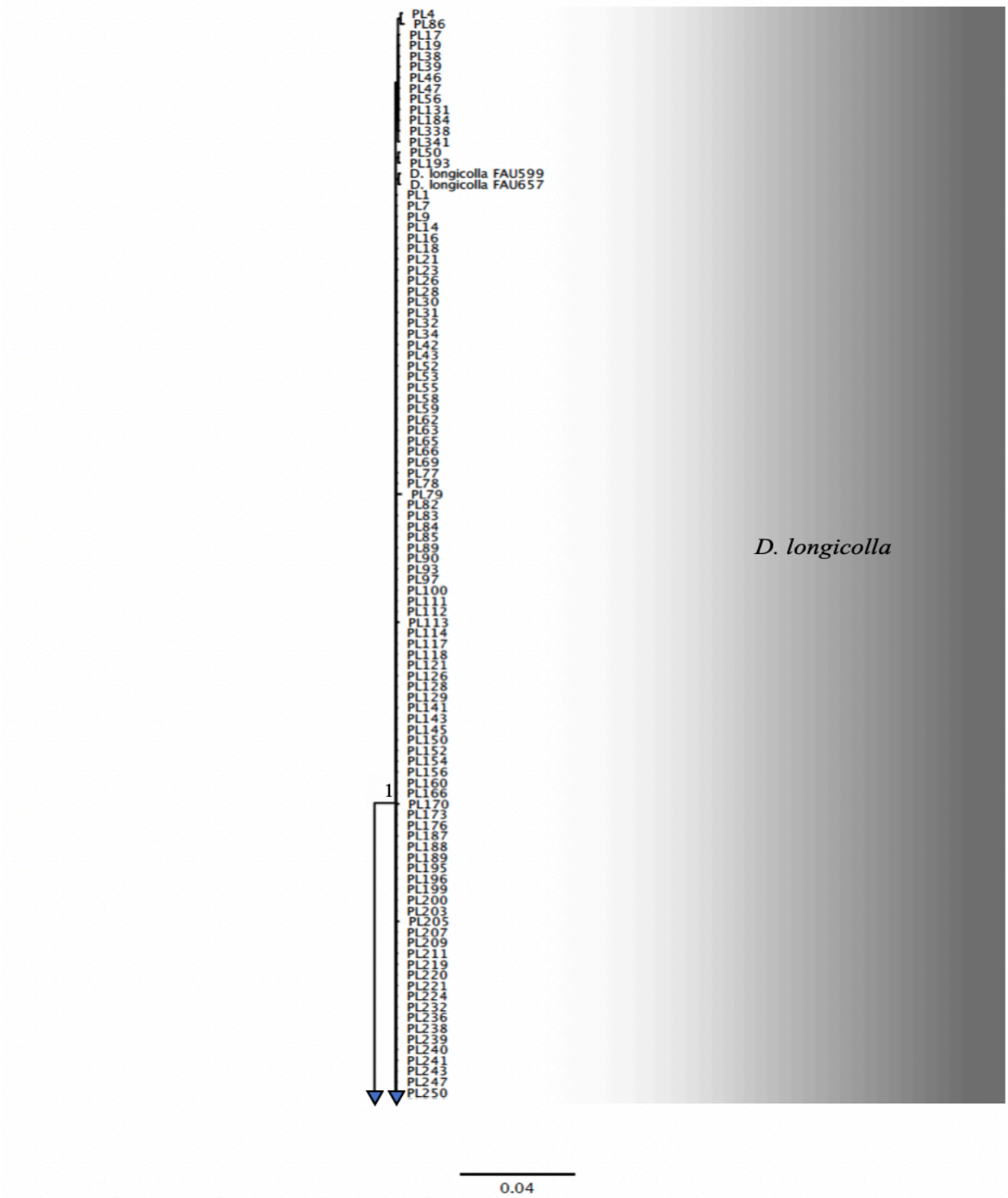


Figure 2. Phylogenetic tree obtained through Bayesian inference (BI) with the concatenated nucleotide sequences of the rDNA ITS region, TEF1-a, TUB2, and CAL loci. The tree includes 184 isolates from in this study and 20 additional fungal isolates. The posterior probability values are displayed above or below each ancestor branch.

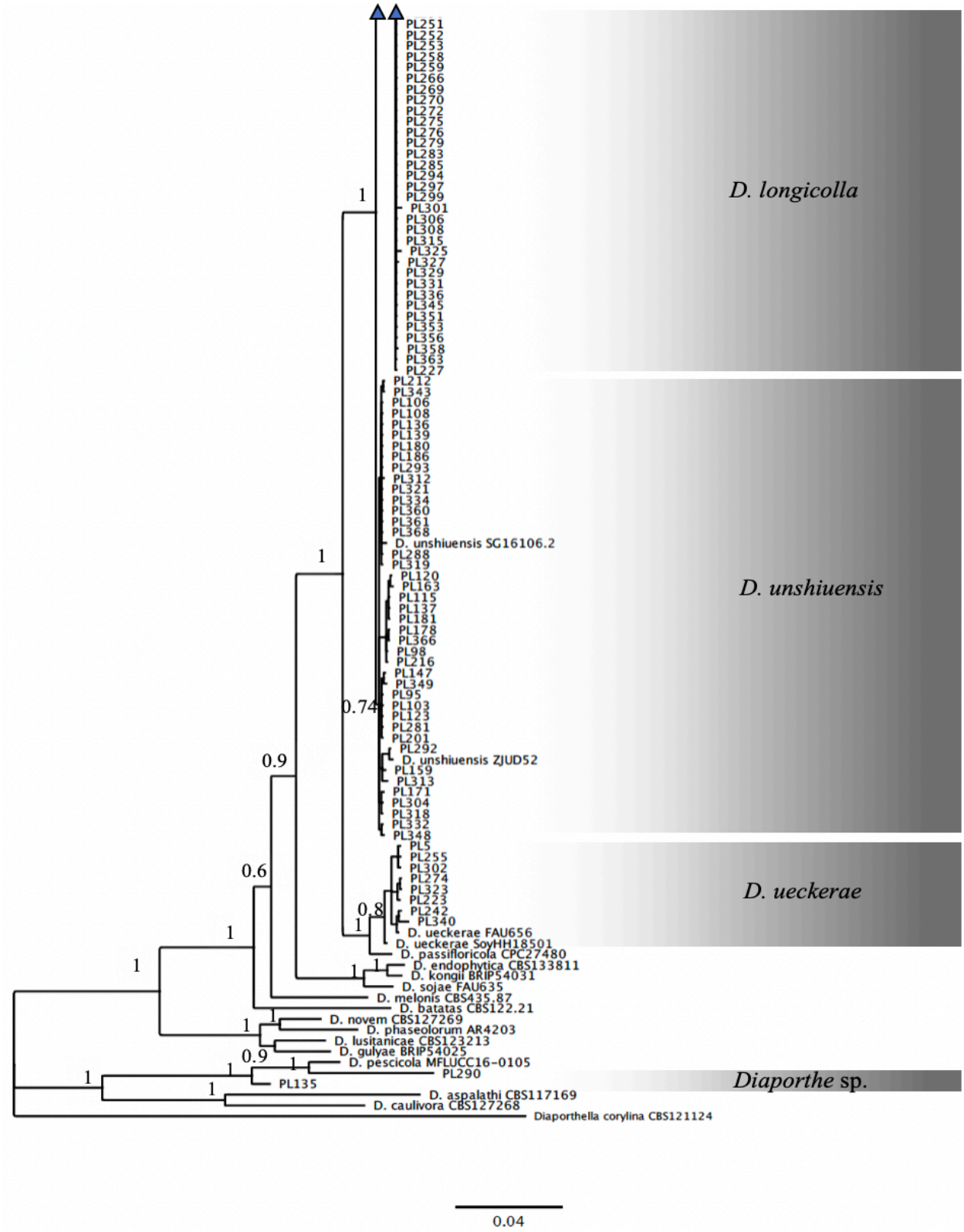


Figure 2. Cont.

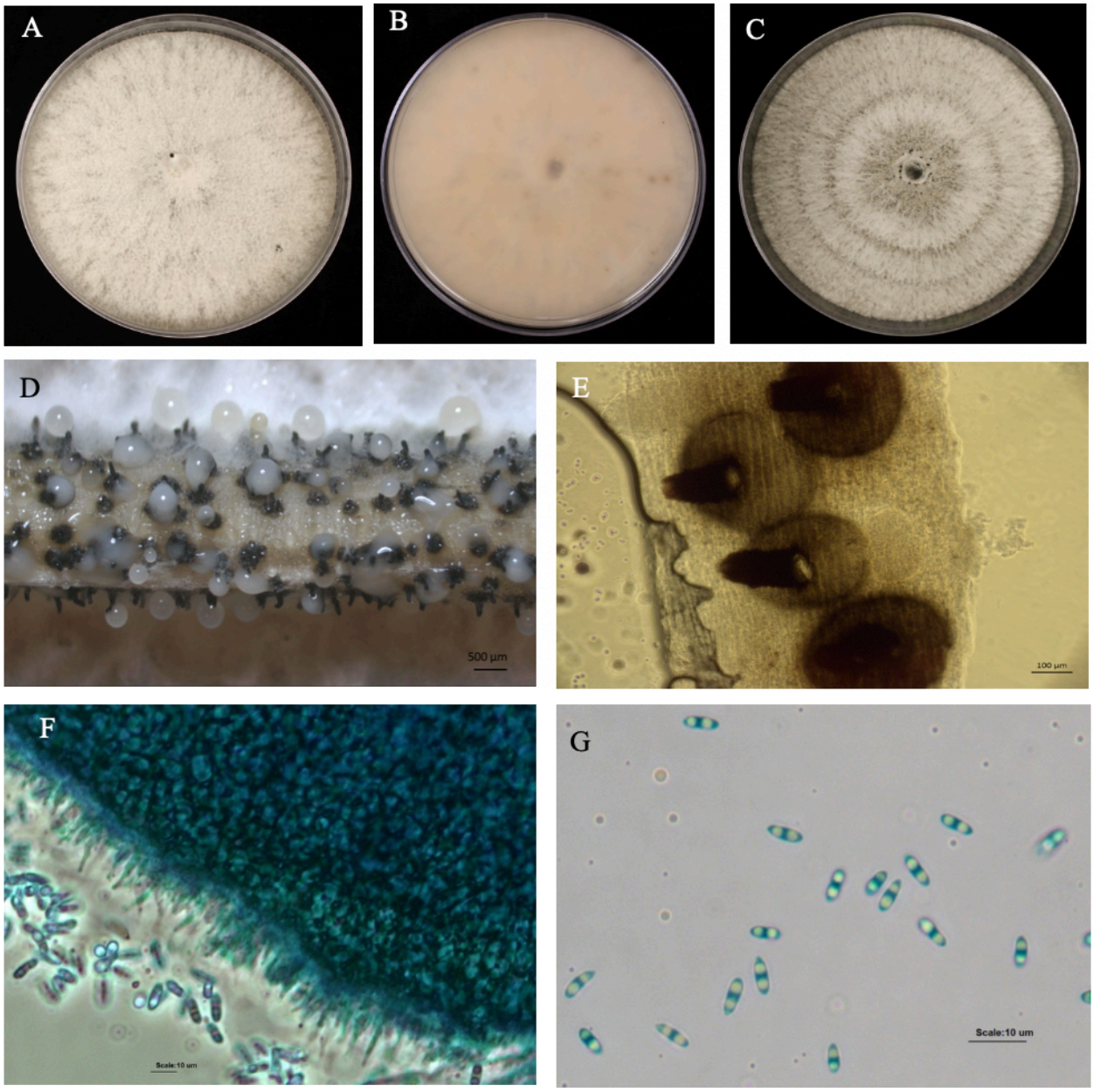


Figure 3. *Diaporthe unshiuensis* (strain PL120). (**A, B**) Culture on PDA after two weeks. (**C**) Culture on OMA after two weeks. (**D**) Conidiomata with conidial mass on soybean stem in culture. (**E**) Pycnidia with long necks. (**F**) Conidiophores. (**G**) Alpha conidia.

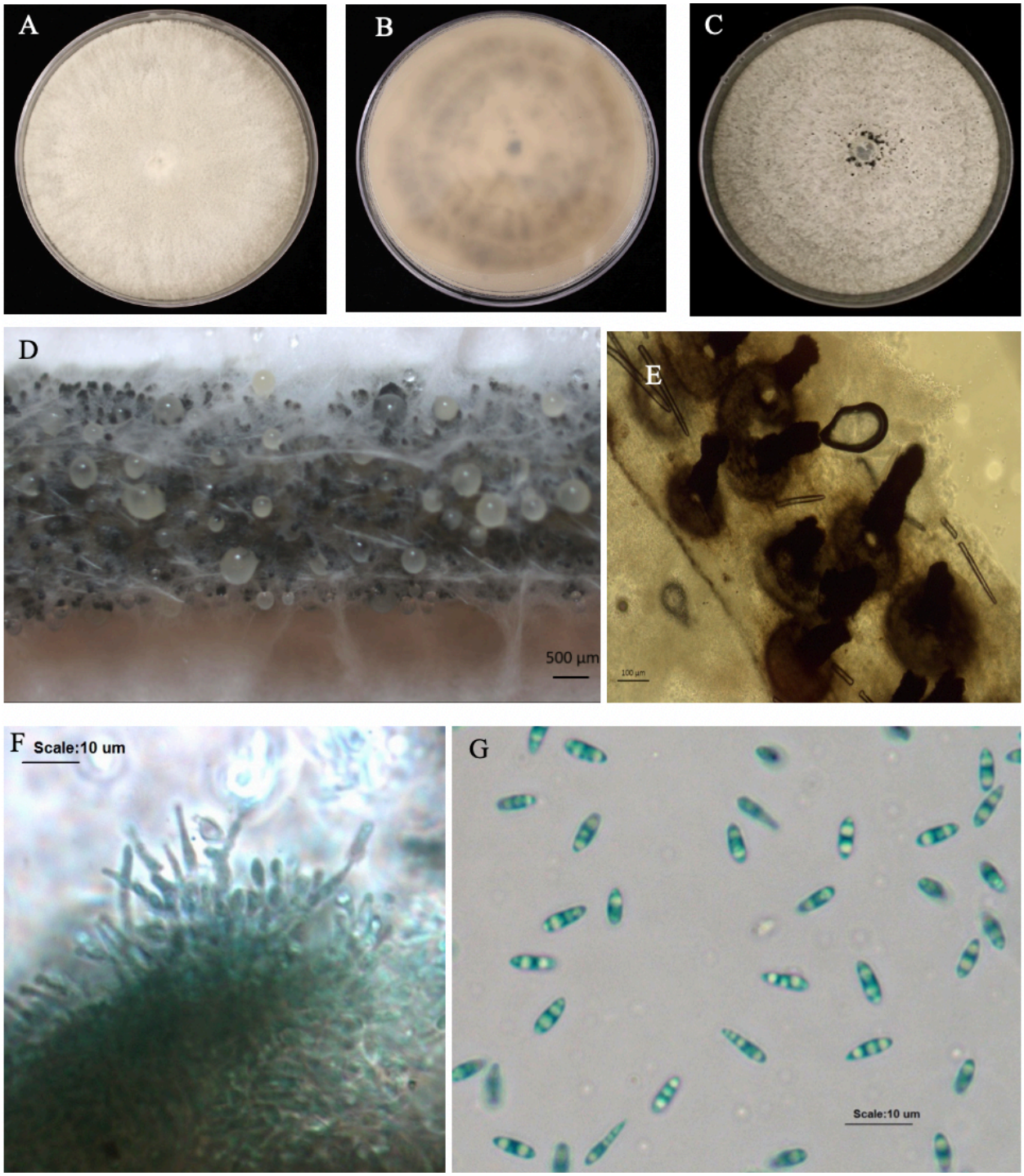


Figure 4. *Diaporthe ueckerae* (strain PL323). **(A, B)** Culture on PDA after two weeks. (**C**) Culture on OMA after two weeks. (**D**) Conidiomata with conidial mass on soybean stem in a culture. (**E**) Pycnidia with long necks. (**F**) Conidiophores. (**G**) Alpha conidia.

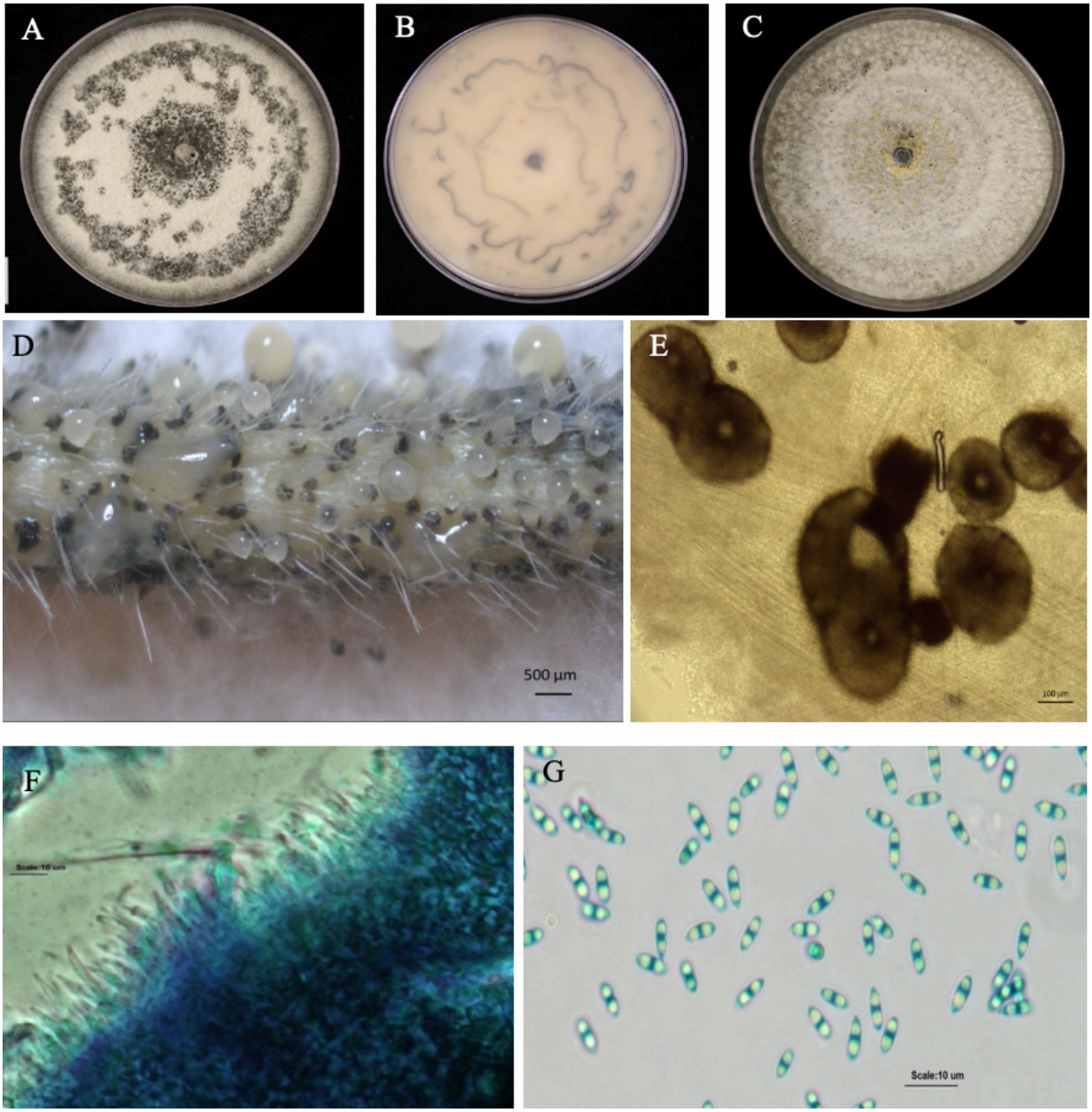


Figure 5. *Diaporthe longicolla* (strain PL82). (**A, B**) Culture on PDA after two weeks. (**C**) Culture on OMA after two weeks. (**D**) Conidiomata with conidial mass on soybean stem in culture. (**E**) Pycnidia with long necks. (**F**) Conidiophores. (**G**) Alpha conidia.

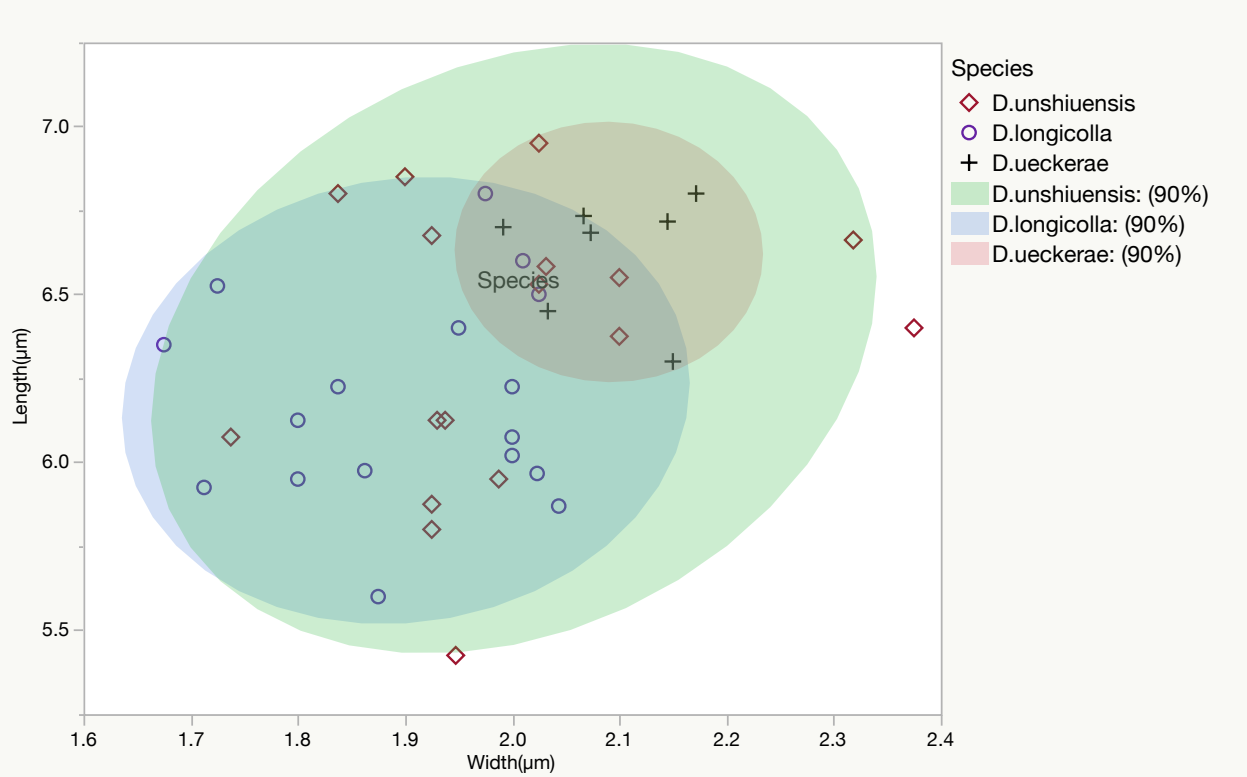


Figure 6. Mean values of length and width of alpha conidia from 17 isolates of *D. unshiuensis*, 17 isolates of *D. longicolla*, and 7 isolates of *D. ueckerae*. Conidia were produced on soybean stems in a 12 h photoperiod at 23 °C within two weeks.

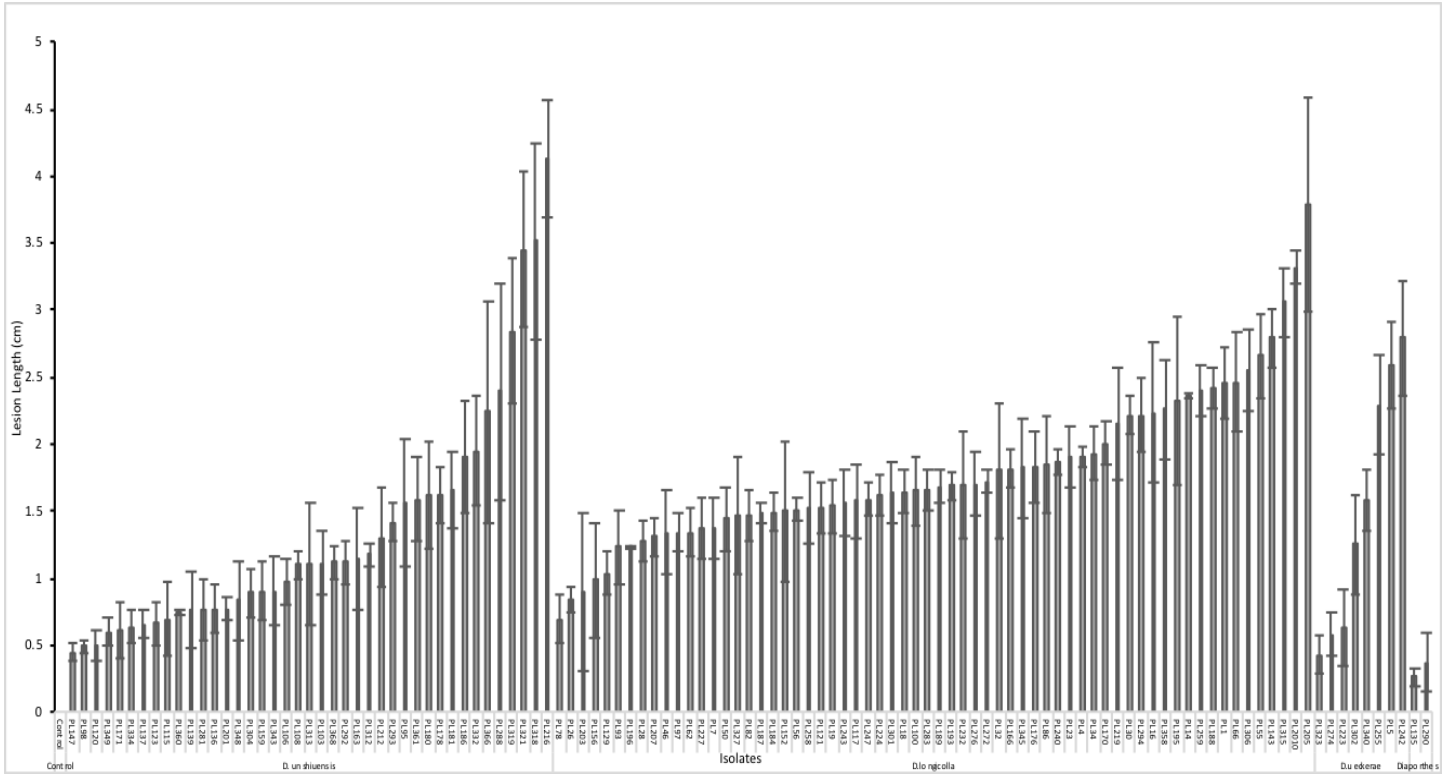


Figure 7. Cut-stem pathogenicity assay of 115 isolates of *Diaporthe* spp. Bars represent standard error of the mean of three replicates. Lesion lengths were measured on stems seven days post inoculation to evaluate pathogenesis.

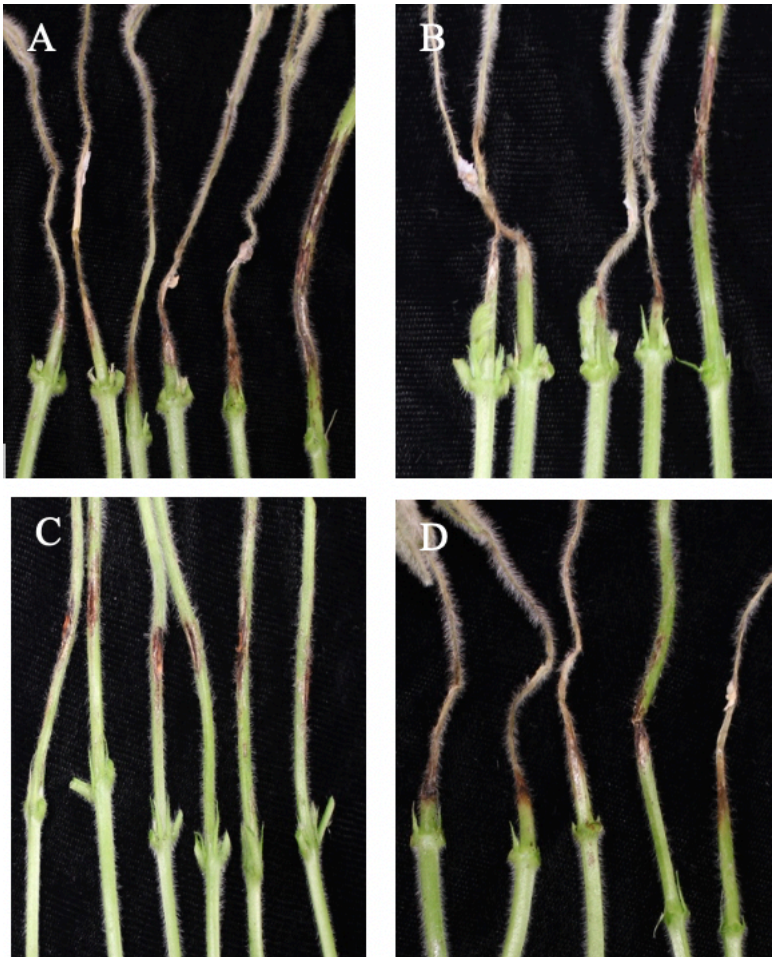


Figure 8. Soybean stem necrosis caused by *Diaporthe* species 10 days post-inoculation with stem-wound inoculation method. Stems were wounded with **(A)** *D. unshiuensis* (strain PL216), **(B)** *D. ueckerae* (strain PL242), **(C)** *Diaporthe* sp. (strain PL135), and **(D)** wild-type strain PL2010 of *D. longicolla.*

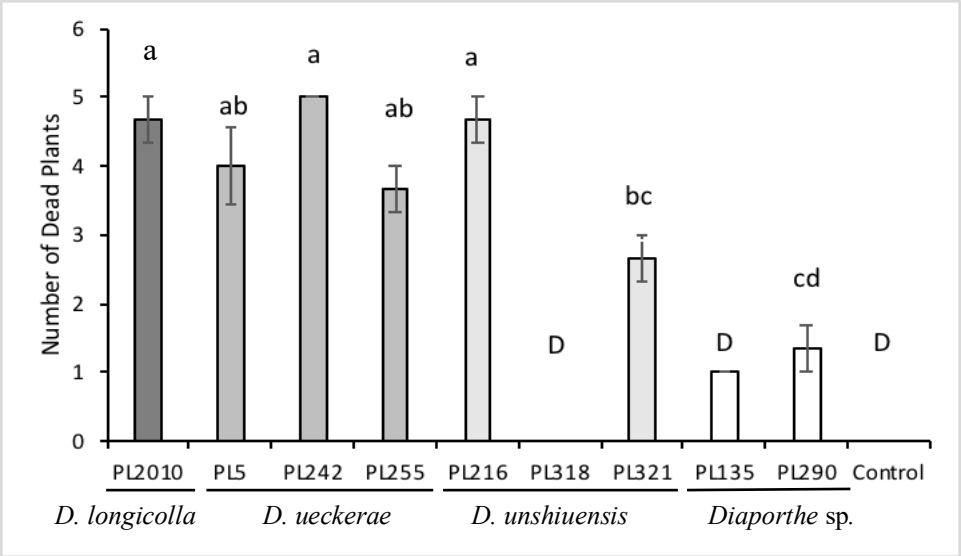


Figure 9. Soybean shoot death (above point of inoculation) induced by nine isolates of *Diaporthe* spp. Data were recorded four days after inoculation. Letters indicate statistically significant differences among average values according to the Tukey test ($P < 0.05$). Each bar represents the average of three replicates, with each replicate comprised of five individual plants.

Table 1. GenBank accession numbers of sequences used in this study.

Species	Isolate	Host	ITS	β - tubulin	Tef1 - α	CAL
D. aspalathi	CBS 117169*	*Aspalathus linearis*	KC343036	KC344004	KC343762	KC343278
D. batatas	CBS 122.21*	*Ipomoea batatas*	KC343040	KC344008	KC343766	KC343282
D. caulivora	CBS 127268*	*Glycine max*	KC343045	KC344013	KC343771	KC343287
D. endophytica	CBS 133811*	*Schinus terebinthifolius*	KC343065	KC343065	KC343791	KC343307
D. gulyae	BRIP 54025*	*Helianthus annuus*	JF431299	N/A [a]	JN645803	N/A [a]
D. kongii	BRIP 54031*	*Helianthus annuus*	JF431301	N/A [a]	JN645797	N/A [a]
D. longicolla	ATCC 60325*	*Glycine max*	KJ590728	KJ610883	KJ590767	N/A [a]
	FAU657	*Cucumis melo*	KJ590727	KJ610882	KJ590766	KJ612123
D. lusitanicae	CBS 123213 *	*Foeniculum vulgare*	KC343137	KC344105	KC343863	KC343379
D. melonis	CBS435.87	*Glycine soja*	KC343141	KC344109	KC343867	KC343383
D. novem	CBS127269	*Glycine max*	KC343155	KC344123	KC343881	KC343397
D. passifloricola	CPC27480*	*Passiflora foetida*	KX228292	KX228387	N/A [a]	N/A [a]
D. pescicola	MFLUCC 16 - 0105*	*Prunus persica*	KU557555	KU557579	KU557623	KU557603
D. phaseolorum	AR4203*	*Phas eolus vulgaris*	KJ590738	KJ610893	KJ590739	KJ612135
D. sojae	FAU635*	*Glycine max*	KJ590719	KJ610875	KJ590762	KJ612116
D. ueckerae	FAU656*	*Cucumis melo*	KJ590726	KJ610881	KJ590747	KJ612122
	SOYHH18501	*Glycine max stem* Myanmar	LC461975	LC461985	LC461991	N/A [a]
D. unshiuensis	ZJUD52*	*Citrus* sp.	KJ490587	KJ490408	KJ490466	N/A [a]
	SG16106.2	*Asparagus kiusianus*	MF185131	MF1950338	MF195045	N/A [a]
Diaporthella corylina	CBS 121124*	*Corylus sp .*	KC343004	KC343972	KC343730	KC343246

a= Sequence not available in GenBank.
* Sequence from type material

Table 2. Distribution and diversity of endophytic *Diaporthe* species associated with soybean stems in Arkansas.

Sampled Sites	*D. longicolla*	*D. unshiuensis*	*D. ueckerae*	*Diaporthe* sp.	SUM	Occurrence %	STDEV	S	E	H	D
Stuttgart	41	0	0	0	41	82	20.5	1	0.0	0.0	0.0
Marianna	27	19	0	1	47	94	13.4	3	0.7	0.77	0.5
Keiser	40	4	4	0	48	96	18.8	3	0.5	0.57	0.3
Rohwer	25	18	4	1	48	96	11.4	4	0.7	1.00	0.6
Average	33.25	10.25	2	0.5	46	92.5	16.02	2.8	0.48	0.58	0.34

SUM = Number of isolates per site
Occurrence % = Percentage of colonized plants (50 plants)
STDEV =Standard Deviation
S = Richness
E = Evenness
H = Shannon`s diversity index
D = Simpson`s diversity index for infinite population

Supplementary Figures and Tables

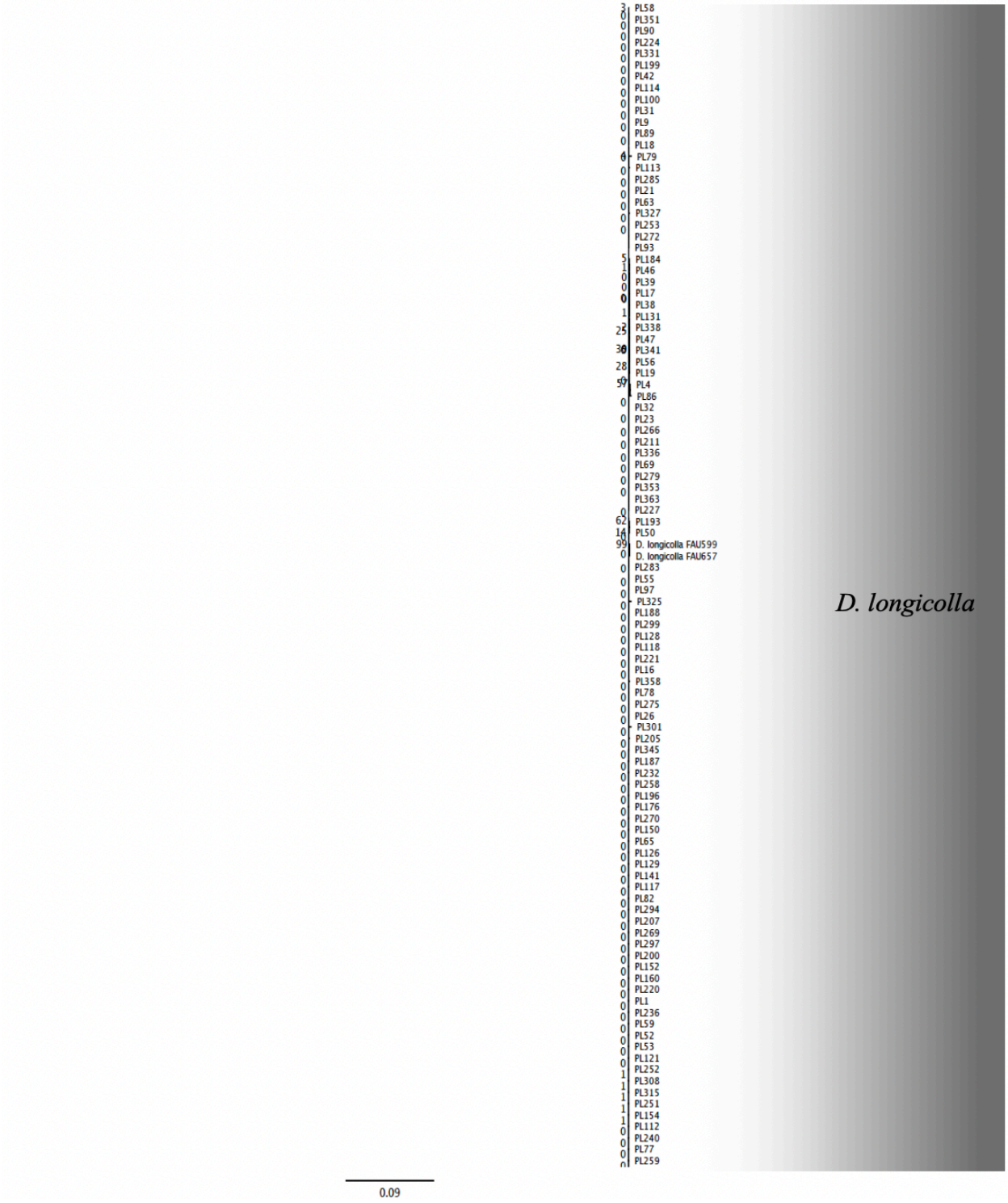


Figure 1S. Phylogenetic tree obtained through maximum likelihood analysis (ML) with the concatenated nucleotide sequence of the rDNA ITS region and the TEF1-α, TUB2 and CAL genes from 184 isolates included in this study and 20 additional isolates (sequences obtained from GenBank). The support probability values are displayed for each ancestor branch.

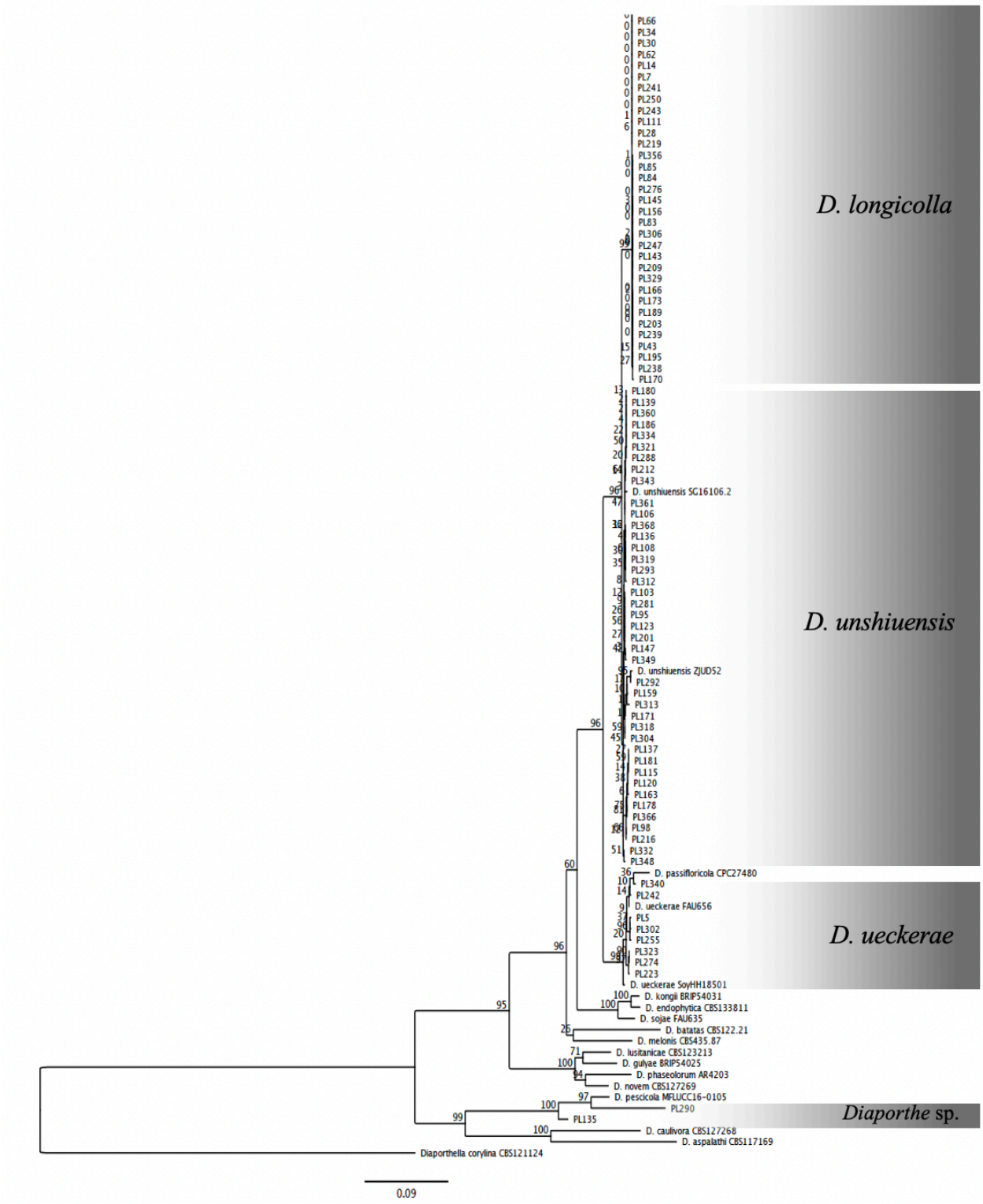


Figure 1S. Cont.

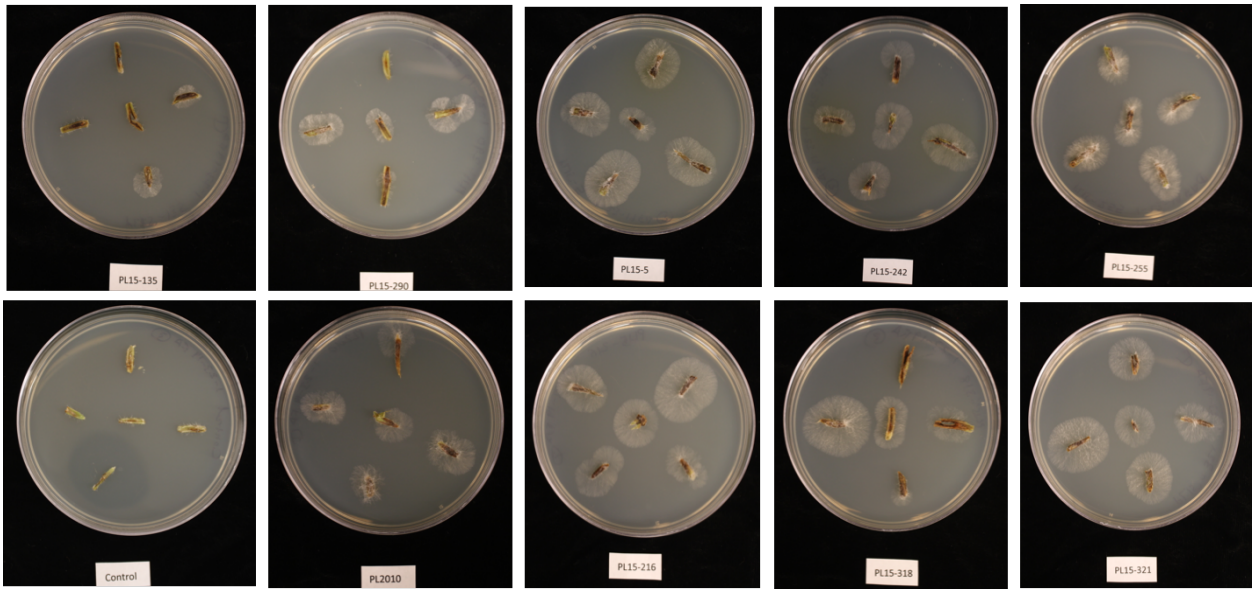


Figure 2S. Re-isolation of *Diaporthe* strains from symptomatic soybean stem tissue.

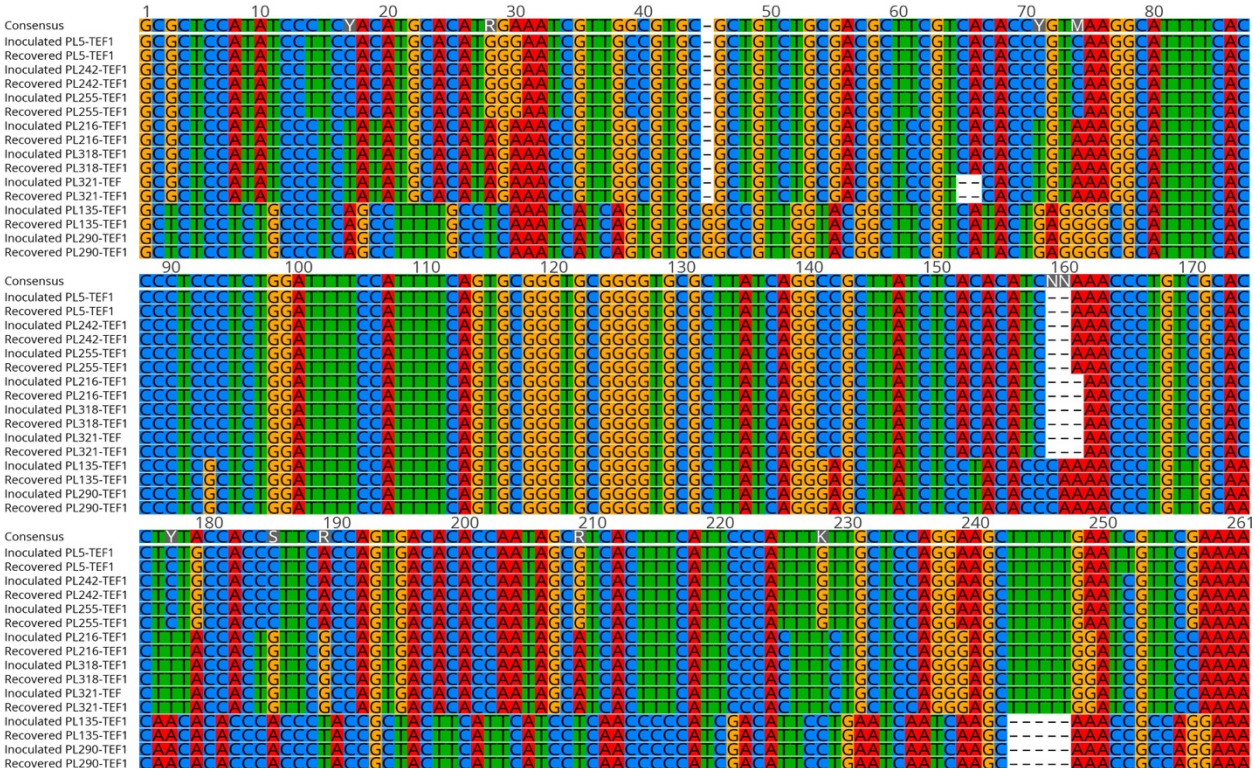


Figure 3S. Alignment of TEF1-α sequences from inoculated and recovered isolates of *D. ueckerae* (strains PL5, PL242, and PL255), *D. unshiuensis* (strains PL216, PL318, and PL321), and *Diaporthe* sp. (strains PL135 and PL290).

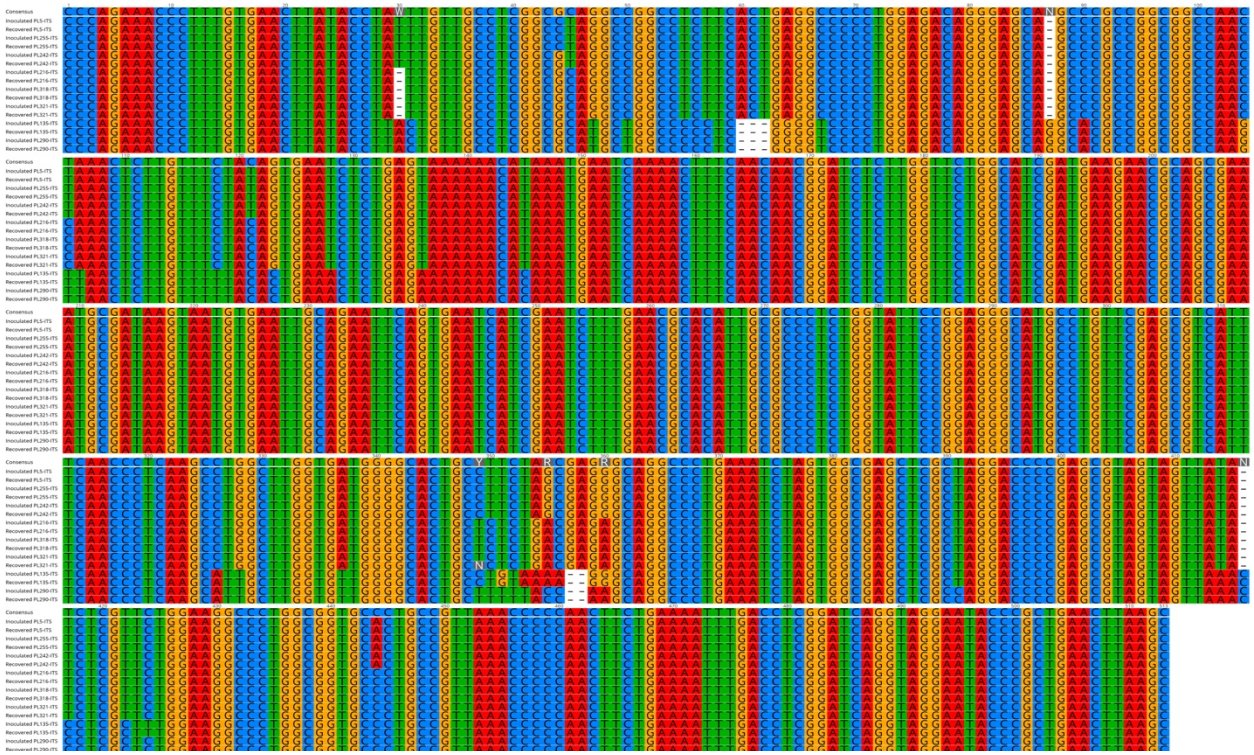


Figure 4S. Alignment of ITS sequences from inoculated and recovered isolates of *D. ueckerae* (strains PL5, PL242, and PL255), *D. unshiuensis* (strains PL216, PL318, and PL321), and *Diaporthe* sp. (strains PL135 and PL290).

Table 1S. Primers used to amplify loci for phylogenetic analysis.

Locus	Primer	Sequence
ITS	ITS1	TCCGTAGGTGAACCTGCGG
	ITS4	TCCTCCGCTTATTGATATGC
TEF1-α	EF1-728F	CATCGAGAAGTTCGAGAAGG
	EF1-986R	TACTTCAAGGAACCCTTACC
TUB2	Bt2a	GGTAACCAAATCGGTGCTGCTTTC
	Bt2b	ACCCTCAGTGTAGTGACCCTTGGC
CAL	CAL563F	GACAAATCACCACCAARGAGC
	CL2A	TTTTTGCATCATGAGTTGGAC
	CL1F	GARTWCAAGGAGGCCTTCTC

Table 2S. GenBank accession numbers of *Diaporthe* species sequenced in this study.

Isolates	Species	ITS	TEF1-α	TUB2	CAL
PL1	*Diaporthe longicolla*	MN586627	MN651137	MN698395	MN725831
PL4	*Diaporthe longicolla*	MN586628	MN651138	MN698396	MN725832
PL5	*Diaporthe ueckerae*	MN586629	MN651139	MN698397	MN725833
PL7	*Diaporthe longicolla*	MN586630	MN651140	MN698398	MN725834
PL9	*Diaporthe longicolla*	MN586631	MN651141	MN698399	MN725835
PL14	*Diaporthe longicolla*	MN586632	MN651142	MN698400	MN725836
PL16	*Diaporthe longicolla*	MN586633	MN651143	MN698401	MN725837
PL17	*Diaporthe longicolla*	MN586634	MN651144	MN698402	MN725838
PL18	*Diaporthe longicolla*	MN586635	MN651145	MN698403	MN725839
PL19	*Diaporthe longicolla*	MN586636	MN651146	MN698404	MN725840
PL21	*Diaporthe longicolla*	MN586637	MN651147	MN698405	MN725841
PL23	*Diaporthe longicolla*	MN586638	MN651148	MN698406	MN725842
PL26	*Diaporthe longicolla*	MN586639	MN651149	MN698407	MN725843
PL28	*Diaporthe longicolla*	MN586640	MN651150	MN698408	MN725844
PL30	*Diaporthe longicolla*	MN586641	MN651151	MN698409	MN725845
PL31	*Diaporthe longicolla*	MN586642	MN651152	MN698410	MN725846
PL32	*Diaporthe longicolla*	MN586643	MN651153	MN698411	MN725847
PL34	*Diaporthe longicolla*	MN586644	MN651154	MN698412	MN725848
PL38	*Diaporthe longicolla*	MN586645	MN651155	MN698413	MN725849
PL39	*Diaporthe longicolla*	MN586646	MN651156	MN698414	MN725850
PL42	*Diaporthe longicolla*	MN586647	MN651157	MN698415	MN725851
PL43	*Diaporthe longicolla*	MN586648	MN651158	MN698416	MN725852
PL46	*Diaporthe longicolla*	MN586649	MN651159	MN698417	MN725853
PL47	*Diaporthe longicolla*	MN586650	MN651160	MN698418	MN725854
PL50	*Diaporthe longicolla*	MN586651	MN651161	MN698419	MN725855
PL52	*Diaporthe longicolla*	MN586652	MN651162	MN698420	MN725856
PL53	*Diaporthe longicolla*	MN586653	MN651163	MN698421	MN725857
PL55	*Diaporthe longicolla*	MN586654	MN651164	MN698422	MN725858
PL56	*Diaporthe longicolla*	MN586655	MN651165	MN698423	MN725859
PL58	*Diaporthe longicolla*	MN586656	MN651166	MN698424	MN725860
PL59	*Diaporthe longicolla*	MN586657	MN651167	MN698425	MN725861
PL62	*Diaporthe longicolla*	MN586658	MN651168	MN698426	MN725862
PL63	*Diaporthe longicolla*	MN586659	MN651169	MN698427	MN725863
PL65	*Diaporthe longicolla*	MN586660	MN651170	MN698428	MN725864
PL66	*Diaporthe longicolla*	MN586661	MN651171	MN698429	MN725865
PL69	*Diaporthe longicolla*	MN586662	MN651172	MN698430	MN725866
PL77	*Diaporthe longicolla*	MN586663	MN651173	MN698431	MN725867
PL78	*Diaporthe longicolla*	MN586664	MN651174	MN698432	MN725868
PL79	*Diaporthe longicolla*	MN586665	MN651175	MN698433	MN725869
PL82	*Diaporthe longicolla*	MN586666	MN651176	MN698434	MN725870
PL83	*Diaporthe longicolla*	MN586667	MN651177	MN698435	MN725871
PL84	*Diaporthe longicolla*	MN586668	MN651178	MN698436	MN725872
PL85	*Diaporthe longicolla*	MN586669	MN651179	MN698437	MN725873
PL86	*Diaporthe longicolla*	MN586670	MN651180	MN698438	MN725874
PL89	*Diaporthe longicolla*	MN586671	MN651181	MN698439	MN725875
PL90	*Diaporthe longicolla*	MN586672	MN651182	MN698440	MN725876
PL93	*Diaporthe longicolla*	MN586673	MN651183	MN698441	MN725877
PL95	*Diaporthe unshiuensis*	MN586674	MN651184	MN698442	MN725878
PL97	*Diaporthe longicolla*	MN586675	MN651185	MN698443	MN725879
PL98	*Diaporthe unshiuensis*	MN586676	MN651186	MN698444	MN725880
PL100	*Diaporthe longicolla*	MN586677	MN651187	MN698445	MN725881

Table 2S. Cont.

Isolates	Species	ITS	TEF1-α	TUB2	CAL
PL103	*Diaporthe unshiuensis*	MN586678	MN651188	MN698446	MN725882
PL106	*Diaporthe unshiuensis*	MN586679	MN651189	MN698447	MN725883
PL108	*Diaporthe unshiuensis*	MN586680	MN651190	MN698448	MN725884
PL111	*Diaporthe longicolla*	MN586681	MN651191	MN698449	MN725885
PL112	*Diaporthe longicolla*	MN586682	MN651192	MN698450	MN725886
PL113	*Diaporthe longicolla*	MN586683	MN651193	MN698451	MN725887
PL114	*Diaporthe longicolla*	MN586684	MN651194	MN698452	MN725888
PL115	*Diaporthe unshiuensis*	MN586685	MN651195	MN698453	MN725889
PL117	*Diaporthe longicolla*	MN586686	MN651196	MN698454	MN725890
PL118	*Diaporthe longicolla*	MN586687	MN651197	MN698455	MN725891
PL120	*Diaporthe unshiuensis*	MN586688	MN651198	MN698456	MN725892
PL121	*Diaporthe longicolla*	MN586689	MN651199	MN698457	MN725893
PL123	*Diaporthe unshiuensis*	MN586690	MN651200	MN698458	MN725894
PL126	*Diaporthe longicolla*	MN586691	MN651201	MN698459	MN725895
PL128	*Diaporthe longicolla*	MN586692	MN651202	MN698460	MN725896
PL129	*Diaporthe longicolla*	MN586693	MN651203	MN698461	MN725897
PL131	*Diaporthe longicolla*	MN586694	MN651204	MN698462	MN725898
PL135	*Diaporthe* sp.	MN586695	MN651205	MN698463	MN725899
PL136	*Diaporthe unshiuensis*	MN586696	MN651206	MN698464	MN725900
PL137	*Diaporthe unshiuensis*	MN586697	MN651207	MN698465	MN725901
PL139	*Diaporthe unshiuensis*	MN586698	MN651208	MN698466	MN725902
PL141	*Diaporthe longicolla*	MN586699	MN651209	MN698467	MN725903
PL143	*Diaporthe longicolla*	MN586700	MN651210	MN698468	MN725904
PL145	*Diaporthe longicolla*	MN586701	MN651211	MN698469	MN725905
PL147	*Diaporthe unshiuensis*	MN586702	MN651212	MN698470	MN725906
PL150	*Diaporthe longicolla*	MN586703	MN651213	MN698471	MN725907
PL152	*Diaporthe longicolla*	MN586704	MN651214	MN698472	MN725908
PL154	*Diaporthe longicolla*	MN586705	MN651215	MN698473	MN725909
PL156	*Diaporthe longicolla*	MN586706	MN651216	MN698474	MN725910
PL159	*Diaporthe unshiuensis*	MN586707	MN651217	MN698475	MN725911
PL160	*Diaporthe longicolla*	MN586708	MN651218	MN698476	MN725912
PL163	*Diaporthe unshiuensis*	MN586709	MN651219	MN698477	MN725913
PL166	*Diaporthe longicolla*	MN586710	MN651220	MN698478	MN725914
PL170	*Diaporthe longicolla*	MN586711	MN651221	MN698479	MN725915
PL171	*Diaporthe unshiuensis*	MN586712	MN651222	MN698480	MN725916
PL173	*Diaporthe longicolla*	MN586713	MN651223	MN698481	MN725917
PL176	*Diaporthe longicolla*	MN586714	MN651224	MN698482	MN725918
PL178	*Diaporthe unshiuensis*	MN586715	MN651225	MN698483	MN725919
PL180	*Diaporthe unshiuensis*	MN586716	MN651226	MN698484	MN725920
PL181	*Diaporthe longicolla*	MN586717	MN651227	MN698485	MN725921
PL184	*Diaporthe longicolla*	MN586718	MN651228	MN698486	MN725922
PL186	*Diaporthe unshiuensis*	MN586719	MN651229	MN698487	MN725923
PL187	*Diaporthe longicolla*	MN586720	MN651230	MN698488	MN725924
PL188	*Diaporthe longicolla*	MN586721	MN651231	MN698489	MN725925
PL189	*Diaporthe longicolla*	MN586722	MN651232	MN698490	MN725926
PL193	*Diaporthe longicolla*	MN586723	MN651233	MN698491	MN725927
PL195	*Diaporthe longicolla*	MN586724	MN651234	MN698492	MN725928
PL196	*Diaporthe longicolla*	MN586725	MN651235	MN698493	MN725929
PL199	*Diaporthe longicolla*	MN586726	MN651236	MN698494	MN725930
PL200	*Diaporthe longicolla*	MN586727	MN651237	MN698495	MN725931
PL201	*Diaporthe unshiuensis*	MN586728	MN651238	MN698496	MN725932
PL203	*Diaporthe longicolla*	MN586729	MN651239	MN698497	MN725933
PL205	*Diaporthe longicolla*	MN586730	MN651240	MN698498	MN725934
PL207	*Diaporthe longicolla*	MN586731	MN651241	MN698499	MN725935
PL209	*Diaporthe longicolla*	MN586732	MN651242	MN698500	MN725936

Table 2S. Cont.

Isolates	Species	ITS	TEF1-α	TUB2	CAL
PL211	*Diaporthe longicolla*	MN586733	MN651243	MN698501	MN725937
PL212	*Diaporthe unshiuensis*	MN586734	MN651244	MN698502	MN725938
PL216	*Diaporthe unshiuensis*	MN586735	MN651245	MN698503	MN725939
PL219	*Diaporthe longicolla*	MN586736	MN651246	MN698504	MN725940
PL220	*Diaporthe longicolla*	MN586737	MN651247	MN698505	MN725941
PL221	*Diaporthe longicolla*	MN586738	MN651248	MN698506	MN725942
PL223	*Diaporthe ueckerae*	MN586739	MN651249	MN698507	MN725943
PL224	*Diaporthe longicolla*	MN586740	MN651250	MN698508	MN725944
PL227	*Diaporthe longicolla*	MN586741	MN651251	MN698509	MN725945
PL232	*Diaporthe longicolla*	MN586742	MN651252	MN698510	MN725946
PL236	*Diaporthe longicolla*	MN586743	MN651253	MN698511	MN725947
PL238	*Diaporthe longicolla*	MN586744	MN651254	MN698512	MN725948
PL239	*Diaporthe longicolla*	MN586745	MN651255	MN698513	MN725949
PL240	*Diaporthe longicolla*	MN586746	MN651256	MN698514	MN725950
PL241	*Diaporthe longicolla*	MN586747	MN651257	MN698515	MN725951
PL242	*Diaporthe ueckerae*	MN586748	MN651258	MN698516	MN725952
PL243	*Diaporthe longicolla*	MN586749	MN651259	MN698517	MN725953
PL247	*Diaporthe longicolla*	MN586750	MN651260	MN698518	MN725954
PL250	*Diaporthe longicolla*	MN586751	MN651261	MN698519	MN725955
PL251	*Diaporthe longicolla*	MN586752	MN651262	MN698520	MN725956
PL252	*Diaporthe longicolla*	MN586753	MN651263	MN698521	MN725957
PL253	*Diaporthe longicolla*	MN586754	MN651264	MN698522	MN725958
PL255	*Diaporthe ueckerae*	MN586755	MN651265	MN698523	MN725959
PL258	*Diaporthe longicolla*	MN586756	MN651266	MN698524	MN725960
PL259	*Diaporthe longicolla*	MN586757	MN651267	MN698525	MN725961
PL266	*Diaporthe longicolla*	MN586758	MN651268	MN698526	MN725962
PL269	*Diaporthe longicolla*	MN586759	MN651269	MN698527	MN725963
PL270	*Diaporthe longicolla*	MN586760	MN651270	MN698528	MN725964
PL272	*Diaporthe longicolla*	MN586761	MN651271	MN698529	MN725965
PL274	*Diaporthe ueckerae*	MN586762	MN651272	MN698530	MN725966
PL275	*Diaporthe longicolla*	MN586763	MN651273	MN698531	MN725967
PL276	*Diaporthe longicolla*	MN586764	MN651274	MN698532	MN725968
PL279	*Diaporthe longicolla*	MN586765	MN651275	MN698533	MN725969
PL281	*Diaporthe unshiuensis*	MN586766	MN651276	MN698534	MN725970
PL283	*Diaporthe longicolla*	MN586767	MN651277	MN698535	MN725971
PL285	*Diaporthe longicolla*	MN586768	MN651278	MN698536	MN725972
PL288	*Diaporthe unshiuensis*	MN586769	MN651279	MN698537	MN725973
PL290	*Diaporthe* sp.	MN586770	MN651280	MN698538	MN725974
PL292	*Diaporthe unshiuensis*	MN586771	MN651281	MN698539	MN725975
PL293	*Diaporthe unshiuensis*	MN586772	MN651282	MN698540	MN725976
PL294	*Diaporthe longicolla*	MN586773	MN651283	MN698541	MN725977
PL297	*Diaporthe longicolla*	MN586774	MN651284	MN698542	MN725978
PL299	*Diaporthe longicolla*	MN586775	MN651285	MN698543	MN725979
PL301	*Diaporthe longicolla*	MN586776	MN651286	MN698544	MN725980
PL302	*Diaporthe ueckerae*	MN586777	MN651287	MN698545	MN725981
PL304	*Diaporthe unshiuensis*	MN586778	MN651288	MN698546	MN725982
PL306	*Diaporthe longicolla*	MN586779	MN651289	MN698547	MN725983
PL308	*Diaporthe longicolla*	MN586780	MN651290	MN698548	MN725984
PL312	*Diaporthe unshiuensis*	MN586781	MN651291	MN698549	MN725985
PL313	*Diaporthe unshiuensis*	MN586782	MN651292	MN698550	MN725986
PL315	*Diaporthe longicolla*	MN586783	MN651293	MN698551	MN725987
PL318	*Diaporthe unshiuensis*	MN586784	MN651294	MN698552	MN725988
PL319	*Diaporthe unshiuensis*	MN586785	MN651295	MN698553	MN725989
PL321	*Diaporthe unshiuensis*	MN586786	MN651296	MN698554	MN725990
PL323	*Diaporthe ueckerae*	MN586787	MN651297	MN698555	MN725991

Table 2S. Cont.

Strain designation	Species	ITS	TEF1-α	TUB2	CAL
PL325	*Diaporthe longicolla*	MN586788	MN651298	MN698556	MN725992
PL327	*Diaporthe longicolla*	MN586789	MN651299	MN698557	MN725993
PL329	*Diaporthe longicolla*	MN586790	MN651300	MN698558	MN725994
PL331	*Diaporthe longicolla*	MN586791	MN651301	MN698559	MN725995
PL332	*Diaporthe unshiuensis*	MN586792	MN651302	MN698560	MN725996
PL334	*Diaporthe unshiuensis*	MN586793	MN651303	MN698561	MN725997
PL336	*Diaporthe longicolla*	MN586794	MN651304	MN698562	MN725998
PL338	*Diaporthe longicolla*	MN586795	MN651305	MN698563	MN725999
PL340	*Diaporthe ueckerae*	MN586796	MN651306	MN698564	MN726000
PL341	*Diaporthe longicolla*	MN586797	MN651307	MN698565	MN726001
PL343	*Diaporthe unshiuensis*	MN586798	MN651308	MN698566	MN726002
PL345	*Diaporthe longicolla*	MN586799	MN651309	MN698567	MN726003
PL348	*Diaporthe unshiuensis*	MN586800	MN651310	MN698568	MN726004
PL349	*Diaporthe unshiuensis*	MN586801	MN651311	MN698569	MN726005
PL351	*Diaporthe longicolla*	MN586802	MN651312	MN698570	MN726006
PL353	*Diaporthe longicolla*	MN586803	MN651313	MN698571	MN726007
PL356	*Diaporthe longicolla*	MN586804	MN651314	MN698572	MN726008
PL358	*Diaporthe longicolla*	MN586805	MN651315	MN698573	MN726009
PL360	*Diaporthe unshiuensis*	MN586806	MN651316	MN698574	MN726010
PL361	*Diaporthe unshiuensis*	MN586807	MN651317	MN698575	MN726011
PL363	*Diaporthe longicolla*	MN586808	MN651318	MN698576	MN726012
PL366	*Diaporthe unshiuensis*	MN586809	MN651319	MN698577	MN726013
PL368	*Diaporthe unshiuensis*	MN586810	MN651320	MN698578	MN726014

Table 3S. Microscopic features of selected *D. unshiuensis* isolates.

Strain designation	Conidia (μm)		Pycnidial diameter (μm)
	L	W	
PL95	(-6)6-7.6 (-7.5)	(-1.5)1.77-2.28(-2.5)	(-255)308-479.7(-510)
PL98	(-6)6.36-7.49(-8)	(-2)1.9-2.3(-2.5)	(-229)343.4-522.9(-561)
PL103	(-5)5.28-6.32(-6.5)	(-1.75)1.74-2.1 (-2.5)	(-255)270.3-609.1(-765)
PL106	(-6)6-7.37 (-8)	(-1.5)1.68-2.17(-2.5)	(-306)365-680(-816)
PL108	(-5)4.85-5.6(-6)	(-1.5)1.52-2(-2)	(-255)371.2-670.7(-765)
PL115	(-5)5.7-6.85(-7)	(-1.75)1.74-2.16(-2.5)	(-255)385.6-670.7(-688)
PL120	(-5)5.942-7.224(-8)	(-2) 1.93-2.136(-2.5)	(-306)356.6-517.5(-637)
PL123	(-6) 6-7.1(-8)	(-2)1.9-2.3(-2.5)	(-255)407.1-602.5(-714)
PL136	(-5)5.8-6.8(-7)	(-1.6)1.8-2.2(-2.5)	(-255)455-697(-765)
PL137	(-5.5)6-7.6(-8)	(-1.5)1.6-2.2(-2.5)	(-350)461-728.9(-750)
PL139	(-5.5)5.7-6.9(-7)	(-1.75)1.8-2.1(-2.5)	(-306)420-574.5(-637)
PL147	(-5.5)6-7.57(-8)	(-1.5)1.6-2(-2)	(-255)314.6-478(-510)
PL159	(-6)6.2-7.5(-8)	(-1.5)1.7-2.1(-2)	(-229)262-376.9(-382)
PL163	(-5.5)5.8-6.4(-6.5)	(-1.5)1.8-2(-2)	(-255)432.5-706.9(-765)
PL171	(-5.5)6-7.4(-9)	(-2)1.9-2.2(-2.5)	(-306)394.8-614.7(765)
PL180	(-5)5.3-6.9(-8)	(_1.5)1.75-2.1(-2.1)	(-255)430.2-653(-688)
PL181	(-5)5.6-7.3(-8)	(-1.5)1.8-2.2(-2.5)	(-229)330-508(-561)
PL186	(-5)5.1-6(-6.5)	(-1.5)1.8-2.2(-2.5)	(-204)216.7-379.7(-561)
PL201	(-5)4.8-6(-6.5)	(-1.5)1.8-2.1(-2)	(-255)267.4-390.1(-459)
PL212	(-6)6-7.9(-9)	(-2)1.9-2.1(-2.5)	((-229)350.1-539.3(-612)
PL216	(-6)5.9-7.5(-8.5)	(-1.5)1.6-2(-2)	(-306)326-546(-637)
PL281	(-6)6-7.4(-8)	(-1.5)1.6-2(-2)	(-255)345-539(-637)
PL288	(-5.5)5.8-8(-10)	(-2)1.95-2.6(-2.5)	(-204)287.5-495(-510)
PL304	(-5)5.4-6.7(-7.5)	(-1.5)1.8-2(-2)	(-204)253.8-372.8(-382)
PL312	(-5)5.3-7(-9)	(-1.5)1.7-2.2(-2.5)	(-306)309.1-473.3(-561)
PL313	(-5)5.6-6.5(-7)	(-1.5)1.6-1.87(-2)	(-255)276.7-416.4(-561)
PL319	(-5)5.7-7(-8)	(-1.5)1.8-2.2(-2.5)	(-280)322.7-508(-663)
PL321	(-5)5.6-7.2(-8)	(-2)1.87-2.8(-3)	(-153)204.3-564.2(-637)
PL334	(-5)5.5-6.6(-7)	(-2)1.9-2.1(-2.5)	(-255)363.4-551.7(-637)
PL343	(-5)5.5-7.3(-9)	(-2)1.3-3.4(-2.5)	(-155)295-464.5(-535)
PL348	(-5)5.7-7(-8)	(-2)1.9-2.3(-2.5)	(-280)335.5-599.7(-765)
PL349	(-4)5.7-7.8(-9)	(-1.5)1.7-2.1(-2.5)	(-204)281.2-371(-382)
PL318	(-5.25)6-7.2(-8)	(-2)1.9-2.7(-2.5)	(-312)412-675(-714)
PL361	(-5)5.4-6.6(-7)	(-1.5)1.8-2(-2)	(-204)221-349.9(-382)
PL366	(-5)5.5-7(-8.5)	(-5)1.6-2(-2)	(-204)253.4-455(-535)
PL368	(-5)5.3-6.4(-7)	(-1.5)1.7-2.1(-2)	(-255)297.8-453.9(-510)

Table 4S. Microscopic features of selected *D. ueckerae* isolates.

Strain designation	Conidia (µm)		Pycnidial diameter (µm)
	L	W	
PL5	(-6)6.2-7.3(-8)	(-2)1.9-2.2(-2.5)	(-153)166-232.7(-281)
PL323	(-6)6.2-7.39 (-8.5)	(-2) 1.97-2.375(-2.5)	(-128)171-249(-280)
PL255	(-6)6.4-7(-7.5)	(-2)1.9-2.3(-2.6)	(-204)268.7-396.3(-382)
PL340	(-6)6.2-7.2(-8)	(-2)1.9-2.25(-2.5)	(-280)319.6-485.8(-561)
PL274	(-6)6-7.4(-80	(-1.5)1.7-2.27(-2.5)	(-255)278.8-488.4(-637)
PL302	(-6)6-6.9(-7)	(-2)1.9-2.2(-2.5)	(-255)289-388(-459)
PL223	(-5)5.7-6.9(-7)	(-1.5)1.8-2.5(-3)	(-255)370-517(-637)

Table 5S. Microscopic features of selected *D. longicolla* isolates.

Strain designation	Conidia (µm)		Pycnidial diameter (µm)
	L	W	
PL1	(-5.5)5.8-7(-7)	(-1.5)1.6-2(-2)	(-127)192-332.8(-459)
PL16	(-5.5)5.7-6.8(-7.5)	(-1.5)1.8-2.2(-2.5)	(-204)252.9-413.3(-510)
PL50	(-5.5)5.8-6.7(-7)	(-1.5)1.6-2(-2)	(-255)311-464(-510)
PL55	(-5)5.9-7.1(-8)	(-1.75)1.8-2.2(-3)	(-255)235-408(-650)
PL78	(-5.5)5.8-7(-8)	(-1.5)1.8-2.1(-2)	(-331)383.8-566.9(-637)
PL82	(-5) 5.53 -6.4 (-7)	(-2) 1.93- 2.12 (-2.5)	(-255)312.9-474.3(-510)
PL93	(-5.5)5.8-6.9(-7.5)	(-1.5)1.5-1.9(-2)	(-357)374.4-584(-765)
PL143	(-5)5.5-6.4(-7)	(-1.5)1.6-1.98(-2)	(-306)343.1-467.1(-510)
PL170	(-5)5.3-6.6(-7)	(-1)1.6-2.1(-2)	(-280)343.4-553.7(-637)
PL184	(-5)5.4-6.4(-7)	(-1.5)1.5-1.94(-2.25)	(-306)342-404.4(-408)
PL188	(-5)5.6-7.4(-8)	(-1.5)1.5-2(-2)	(-280)270-422.8(-637)
PL205	(-5)5-6.1(-6.5)	(-1.5)1.7-2(-2)	(-306)305-459.5(-561)
PL207	(-5)5.6-7.56(-9.5)	(-1.9)1.9-2.1(-2.5)	(-146)203.7-328(-357)
PL332	(-6)6.3-7.3(-8)	(-1.5)1.9-2(-2)	(-331)422-637(-765)

Table 6S. The names of *Diaporthe* isolates deposited in -80 freezer inventory and the alternative names in the dissertation. Only the constant part (15-) in names of all isolates were removed to allow the software to analyze the data.

Isolate Name in -80 freezer inventory	Alternative strain name (in the Dissertation)	Isolate Name in -80 freezer inventory	Alternative strain name (in the Dissertation)	Isolate Name in -80 freezer inventory	Alternative strain name (in the Dissertation)	Isolate Name in -80 freezer inventory	Alternative strain name (in the Dissertation)
PL15-1	PL1	PL15-93	PL93	PL15-186	PL186	PL15-279	PL279
PL15-4	PL4	PL15-95	PL95	PL15-187	PL187	PL15-281	PL281
PL15-5	PL5	PL15-97	PL97	PL15-188	PL188	PL15-283	PL283
PL15-7	PL7	PL15-98	PL98	PL15-189	PL189	PL15-285	PL285
PL15-9	PL9	PL15-100	PL100	PL15-193	PL193	PL15-288	PL288
PL15-14	PL14	PL15-103	PL103	PL15-195	PL195	PL15-290	PL290
PL15-16	PL16	PL15-106	PL106	PL15-196	PL196	PL15-292	PL292
PL15-17	PL17	PL15-108	PL108	PL15-199	PL199	PL15-293	PL293
PL15-18	PL18	PL15-111	PL111	PL15-200	PL200	PL15-294	PL294
PL15-19	PL19	PL15-112	PL112	PL15-201	PL201	PL15-297	PL297
PL15-21	PL21	PL15-113	PL113	PL15-203	PL203	PL15-299	PL299
PL15-23	PL23	PL15-114	PL114	PL15-205	PL205	PL15-301	PL301
PL15-26	PL26	PL15-115	PL115	PL15-207	PL207	PL15-302	PL302
PL15-28	PL28	PL15-117	PL117	PL15-209	PL209	PL15-304	PL304
PL15-30	PL30	PL15-118	PL118	PL15-211	PL211	PL15-306	PL306
PL15-31	PL31	PL15-120	PL120	PL15-212	PL212	PL15-308	PL308
PL15-32	PL32	PL15-121	PL121	PL15-216	PL216	PL15-312	PL312
PL15-34	PL34	PL15-123	PL123	PL15-219	PL219	PL15-313	PL313
PL15-38	PL38	PL15-126	PL126	PL15-220	PL220	PL15-315	PL315
PL15-39	PL39	PL15-128	PL128	PL15-221	PL221	PL15-318	PL318
PL15-42	PL42	PL15-129	PL129	PL15-223	PL223	PL15-319	PL319
PL15-43	PL43	PL15-131	PL131	PL15-224	PL224	PL15-321	PL321
PL15-46	PL46	PL15-135	PL135	PL15-227	PL227	PL15-323	PL323
PL15-47	PL47	PL15-136	PL136	PL15-232	PL232	PL15-325	PL325
PL15-50	PL50	PL15-137	PL137	PL15-236	PL236	PL15-327	PL327
PL15-52	PL52	PL15-139	PL139	PL15-238	PL238	PL15-329	PL329
PL15-53	PL53	PL15-141	PL141	PL15-239	PL239	PL15-331	PL331
PL15-55	PL55	PL15-143	PL143	PL15-240	PL240	PL15-332	PL332
PL15-56	PL56	PL15-145	PL145	PL15-241	PL241	PL15-334	PL334
PL15-58	PL58	PL15-147	PL147	PL15-242	PL242	PL15-336	PL336
PL15-59	PL59	PL15-150	PL150	PL15-243	PL243	PL15-338	PL338
PL15-62	PL62	PL15-152	PL152	PL15-247	PL247	PL15-340	PL340
PL15-63	PL63	PL15-154	PL154	PL15-250	PL250	PL15-341	PL341
PL15-65	PL65	PL15-156	PL156	PL15-251	PL251	PL15-343	PL343
PL15-66	PL66	PL15-159	PL159	PL15-252	PL252	PL15-345	PL345
PL15-69	PL69	PL15-160	PL160	PL15-253	PL253	PL15-348	PL348
PL15-77	PL77	PL15-163	PL163	PL15-255	PL255	PL15-349	PL349
PL15-78	PL78	PL15-166	PL166	PL15-258	PL258	PL15-351	PL351
PL15-79	PL79	PL15-170	PL170	PL15-259	PL259	PL15-353	PL353
PL15-82	PL82	PL15-171	PL171	PL15-266	PL266	PL15-356	PL356
PL15-83	PL83	PL15-173	PL173	PL15-269	PL269	PL15-358	PL358
PL15-84	PL84	PL15-176	PL176	PL15-270	PL270	PL15-360	PL360
PL15-85	PL85	PL15-178	PL178	PL15-272	PL272	PL15-361	PL361
PL15-86	PL86	PL15-180	PL180	PL15-274	PL274	PL15-363	PL363
PL15-89	PL89	PL15-181	PL181	PL15-275	PL275	PL15-366	PL366
PL15-90	PL90	PL15-184	PL184	PL15-276	PL276	PL15-368	PL368

Chapter 3: Genetic Diversity and Identification of *Diaporthe* Species on Soybean Utilizing Genotyping-By-Sequencing of Microsatellites

Abstract

Diaporthe species are associated with important diseases of soybean in Arkansas, including stem canker, pod and stem blight, and Phomopsis seed decay. To assess the population structure of *D. longicolla*, *D. unshiuensis,* and *D. ueckerae* in Arkansas, 90 isolates of these three species were analyzed via genotyping-by-sequencing (GBS) with microsatellite markers (simple sequence repeats, SSRs). With GBS data, genotypic diversity, source of variation, species identification, geographic trend, and genetic isolation were assessed. Eight SSRs with 55 alleles distinguished more than 90% of the unique multilocus genotypes (MLGs) present among the collection of isolates. These markers successfully discriminated isolates of *D. longicolla, D. ueckerae,* and *D. unshiuensis* as clusters via genetic distance information. Furthermore, one microsatellite [C11((TGCC)6)] discriminated the three *Diaporthe* species with five species-specific alleles. Genotypic diversity indices revealed that populations of *Diaporthe* species including *D. longicolla* and *D. unshiuensis* were highly diverse at all sites. However, *D. ueckerae* showed the lowest diversity, possibly due to the low number of individuals analyzed. Furthermore, although *D. ueckerae* and *D. unshiuensis* were not isolated from some of the locations sampled, genotypes of these species were not genetically clustered based on sampling site. Linkage disequilibrium indices indicated that populations of *D. longicolla* are largely clonal in Arkansas, whereas populations of *D. unshiuensis* are likely to be undergoing sexual reproduction. PCR diagnosis of mating-type genes showed that all the examined isolates of the three *Diaporthe* species possess both mating-type genes. Furthermore, vegetative incompatibility between these species was observed by the formation of pigmented barrage reaction lines

between expanding colonies along the zone of mycelial contact during the vegetative compatibility tests. These findings substantially expanded the existing knowledge about genotypic diversity and population structure among *Diaporthe* species associated with soybean in Arkansas and could potentially inform future efforts to develop novel disease management strategies.

Introduction

Diseases of soybean caused by *Diaporthe* species, including stem canker, stem and pod blight, and Phomopsis seed decay, collectively cause substantial yield losses in Arkansas and other U.S. states (Allen et al., 2019). One of the *Diaporthe* species most commonly associated with soybean is *Diaporthe longicolla* (syn. *Phomopsis longicolla*) (Li, 2011). To date, little is known about the genetic basis of pathogenesis in *D. longicolla*. Population genetic analyses of fungal pathogens provide robust tools to investigate how pathogens emerge and adapt (Grunwald et al., 2017). Thus, estimating genotypic diversity, which is a key component of defining population structure among plant pathogens (Grunwald et al., 2003), could help dissect pathogen dynamics in *D. longicolla* and related species.

To thoroughly investigate genetic diversity among populations of *D. longicolla* and closely related species on soybean, robust molecular markers and genotyping techniques are required. For example, microsatellites (simple sequence repeats; SSRs) utilize the hypervariability of DNA regions comprised of multiple tandem repeats of di-, tri- or multiple nucleotides (Xu, 2006). This hypervariability mainly results from unequal crossing-over during meiosis or strand slippage during DNA replication (Xu, 2006). Microsatellites are trustworthy markers for the analysis of genetic diversity (Milgroom, 2015), and, in some cases, taxonomic resolution of closely related fungal species (del Castillo-Munera et al., 2013). Microsatellites can

be detected via traditional gel electrophoresis, capillary electrophoresis with labeled primers, and recently with next-generation DNA sequencing (Stewart et al., 2011; Darby et al., 2016).

Next-generation DNA sequencing (NGS) platforms have revolutionized analyses of genetic variation and population genetics by providing rapid, low-cost sequencing of genetic markers in large populations. NGS applications for diversity analyses include genotyping by sequencing (GBS) for genome-wide identification of single nucleotide polymorphisms (SNPs) (Leboldus et al., 2015), simple sequence repeats (SSRs) (Vartia et al., 2016), and inter simple sequence repeats (ISSR) (Suyama & Matsuki, 2015). To sequence markers via NGS, highly reduced representation libraries can be constructed via PCR without restriction enzyme digestion, which allows simultaneous sequencing of multiple markers from many different individuals (Suyama & Matsuki, 2015). NGS can also be applied to analyses of genetic population structure via GBS of microsatellite loci (Darby et al., 2016). GBS of microsatellites can mitigate technical issues associated with gel and capillary electrophoresis-based analyses, including amplicon size homoplasy, missing underlying sequence data, inter-laboratory calibration, and inherently laborious genotyping (Delmotte et al., 2001; Pasqualotto et al., 2007). Additionally, NGS can help avoid some defects of GBS with SNPs, such as the requirement for high DNA quality and potential ascertainment bias (Helyar et al., 2011; Kuhner et al., 2000; Nielsen, 2000).

Because the population structure of *D. longicolla* and closely related species has not been investigated in Arkansas, important questions remain unaddressed regarding genotypic diversity, genetic variation among species, geographical trends regarding species distribution, genetic isolation between closely related species, and distribution of mating-type loci. To answer these

questions, NGS of microsatellites was utilized to assess genetic diversity among populations of *D. longicolla*, *D. unshiuensis*, and *D. ueckerae* at four locations in Arkansas.

Materials and Methods

Isolates of *Diaporthe* species

A sub-set of 90 *Diaporthe* isolates were selected arbitrarily to represent approximately 50% of the isolates collected from Stuttgart, Marianna, Keiser, and Rohwer in Arkansas (as described in Chapter 2). All 90 of these isolates (20-24 from each of the four sites) were previously identified to the species level in Chapter 2 (Table 1S).

Extraction of SSRs, primer design and identification of polymorphic SSRs

A total of 48 SSRs and their corresponding forward and reverse primers were extracted from the draft genome of *D. longicolla* strain MSPL 10-6 (GenBank accession AYRD00000000; (Li et al., 2015). SSRs were identified with Websat (Martins et al., 2009). To check for multi-allelic SSRs among the 48 markers, 8 isolates of *D. longicolla* were used. SSR loci were amplified in 25 µL PCR reactions containing 20 ng of DNA, 1 µl of 10 mM dNTPs, 30 nM of forward and reverse primers, 2.5 µl of 25 mM $MgCl_2$ and 0.15 U of Taq polymerase. The PCR conditions used were 94 °C denaturation for 5 min, followed by 30 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 30 s and extension at 72 °C for 7 min. PCR products were analyzed via electrophoresis in 3% agarose gels stained with GelRed® Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA, USA). Of the 48 markers, 10 were utilized in subsequent analyses (Table 1S).

DNA extraction, preparing libraries, and sequencing SSRs

Genomic DNA was extracted with a CTAB protocol as described by Leslie and Summerell (2006). The construction of sequencing libraries required two rounds of PCR amplification. In

the first round of PCR, each SSR was amplified in an individual reaction for each isolate. Forward and reverse primers for each SSR contained a 3 bp anchor for the second-round PCR primers, a 5′ tail (14 bp) as an annealing region for the second-round PCR primers, and SSR-specific primer sequences (Table 3S). Amplifications were performed in 25 µl reactions containing 1 µl 10 mM dNTPs, 5 µl PCR buffer (5x), 3 µl DNA template (10 ng /µl), 1.5 µl 10mM primers, 14.2 µl H_2O and 0.3 µl Taq DNA polymerase. Amplification conditions included denaturation (94 °C for 2 min), followed by 40 amplification cycles (30 s at 94 °C, 30 s at 54-57 °C and 30 s at 72 °C), and a final extension cycle (72 °C for 5 min). Reaction products were evaluated in 1% agarose gels via electrophoresis and visualized with GelRed® Nucleic Acid Gel Stain. To create templates for the second round of PCR, the 10 SSR amplicons from each isolate were pooled (4 µl of each amplicon) and diluted 50x with distilled sterile water. The second round of PCR, a tailed PCR utilizing common and indexed primers with a 10-12 bp index for each isolate (Table 4S and 5S), was performed in a 25 µl reactions containing 1 µl 10 mM dNTPs, 5 µl PCR buffer (5x), 3 µl DNA template (10 ng /1 µl), 4 µl 10 mM primers, 11.75 µl H_2O, and 0.25 µl Taq DNA polymerase. Amplification conditions included denaturation (94 °C for 4 min), followed by 38 amplification cycles (30 s at 94 °C, 30 s at 54 °C and 30 s at 72 °C), and a final extension cycle (72 °C for 5 min). PCR products were purified with Agencourt AMPure XP beads (Beckman Coulter, IN, USA) in a ratio of 1:1.8 (v: v) PCR product: beads, following the manufacturer's instructions.

After estimating the concentration of each purified second-round PCR product with a Nanodrop1000 spectrophotometer (Thermo Scientific, MA, USA), products were pooled in equimolar concentrations to create a final, mixed library (Suyama & Matsuki, 2015). The final, mixed library was qualified and validated with an Agilent Tapestation 2200 DIK with D1000

ScreenTape® (Agilent Technologies, Santa Clara, CA) and sequenced on one Ion 314™ chip kit V2 with the Ion Torrent (PGM) sequencing platform (Life Technologies, Carlsbad, USA) according to the protocol of the manufacturer.

Sequencing data and allele calling analyses

DNA sequence data was accessed from the Ion Torrent server with indexed sequences of isolates via barcoded adaptor sequences (Table 1S). Reads were classified based on the forward primer sequences by specifying the parameter *-g* (5' adapter) with Cutadapt v1.14 (Martin, 2011). After demultiplexing and filtering the data, reads smaller than 100 bp were removed with the script *reformat.sh* within the BBMap suite v35.92 (http://bbtools.jgi.doe.gov). The reads of each locus of each isolate were merged to one read as a majority consensus with Geneious version 9.1.8 (Kearse et al., 2012). For some missing data, amplicons were sequence via Sanger sequencing (Genewiz Inc., South Plainfield, NJ, USA). Alleles of each haploid microsatellite locus were coded with the number of repeats manually counted from microsatellite sequences and converted to GeneAlEx data format within Excel as numeric data.

Analysis of microsatellite data

GeneAlEx files were converted to a *.CSV format that was then imported into poppr version 2.8.2 (Kamvar et al., 2014) in R (version 3.5.3) (R Core Team, 2019). Genotype accumulation curves were calculated by determining the minimum number of loci necessary to discriminate between isolates in a population via counting the number of multilocus genotypes observed. Furthermore, the indices of population genetic diversity and structures were calculated in poppr, including Shannon-Weiner diversity (Shannon & Weaver, 1998), Stoddart and Tylor's index, G (Stoddart & Taylor, 1988), Simpson's Index, lambda (Simpson, 1949), Nei's unbiased gene

diversity (Nei, 1978), Evenness, E5 (Grunwald et al., 2003) and the number of expected multi-locus genotypes (eMLG) in a rarefied sample size (n).

Poppr version 2.8.2 was also utilized to construct the minimum spanning network (Csardi & Nepusz, 2006) and UPGMA tree (Schliep, 2011) using genetic distance to represent relatedness of genotypes or isolates of *Diaporthe* species as clusters. Furthermore, the presence and absence of linkage disequilibrium were assessed by calculating the index of association (Ia) (Brown et al., 1980; Smith et al., 1993) and the standard index of association (rbarD). To calculate the degree of genetic variability between and within population components, an analysis of molecular variance (AMOVA) was performed within poppr version 2.8.2. All analyses within the poppr package were performed with and without clone correction and the data were presented with clone correction**.**

Vegetative incompatibility tests

Vegetative incompatibility tests between and within isolates of *Diaporthe* species were performed based on the formation of a barrage-zone (Guillin et al., 2014). Fifteen isolates of *Diaporthe* species (five isolates each of *D. longicolla*, *D. unshiuensis*, and *D. ueckerae*) were tested against other species and isolates of the same species. Fungi were plated as three mycelial plugs (5 mm) obtained from 7-day-old cultures in each 9 cm PDA plate with two replicates. Plates were incubated in darkness for 7 days at 20 °C followed by 7 days at 25 °C. After 14 days, barrage zones were recorded, photographed, and microscopically documented (Guillin et al., 2014).

Mating-type diagnosis

To identify mating-type genes as described by Santos et al. (2010), primers MAT1-1-1FW and MAT1-1-1RV were used to amplify part of the a1 box from MAT1-1-1; and primers MAT1-2-1FW and MAT1-2-1RV were used to amplify part of the HMG domain from MAT1-2-1 (Table 6S). For PCR, reactions (25 µl) contained 5 µl PCR buffer (5x), 1 µl 10 mM dNTPs, 3 µl DNA template, 2 µl of each primer, 11.7 µl H_2O and 0.3 µl of Taq DNA polymerase was used for amplifying part of the MAT1-1-1 and MAT1-2-1 genes. Amplification conditions were 1 cycle of 5 min at 95 ºC; 40 cycles of 30 s at 94 ºC, 30 s (at 50 ºC for MAT1-1-1 or 56 ºC for MAT1-2-1), and 1 min at 72 ºC; and a final step of 10 min at 72 ºC (Santos et al., 2010). All PCR products were evaluated on 1% agarose gels stained with GelRed® Nucleic Acid Gel Stain (Biotium, Inc., Hayward, CA, USA) and visualized with a Gel Doc imaging system (Bio-Rad Laboratories).

Results

Microsatellite selection

A total of 48 microsatellites (SSRs) were identified with at least six repeats per motif, consisting of 37.5% dinucleotide repeats, 35.4% trinucleotide repeats, 18.75%, tetranucleotide repeats, and 4.16% pentanucleotide and hexanucleotide repeats. Of these, 10 SSRs could be scored as a single, polymorphic band for the eight isolates initially selected to evaluate polymorphism (data not shown). Five of these ten SSRs were dinucleotide repeats, four were tetranucleotide repeats, and one was a hexanucleotide repeat (Table S1).

Sequencing output, coverage, rarefication, and allele calling

Sequencing output was 44,503,446 bp representing 316,176 reads. After successfully classifying, trimming, and removing reads less than 100 bp, the total sequencing output was

34,391,525 bp representing 156,092 reads. Read lengths were 100 – 431 bp with an average length of 220 bp and a median length 220 bp. Although two SSRs (B3 and B11) were removed from the analysis due to missing data (Figure 1S), the genotypic accumulation curve of the remaining eight markers indicated that approximately 100% of MLGs presented in the sample could be differentiated among the 90 individuals evaluated. Furthermore, for the subpopulations of *D. longicolla* and *D. unshiuensis*, the eight microsatellite loci revealed a high percent (~100%) of MLGs present in the individuals of these species. These results indicated that the number of SSRs and individuals were sufficient to study *Diaporthe* populations (Figure 1). Additionally, all eight microsatellite loci were polymorphic, with the number of alleles per locus ranging from 3 to 17 (Table 1).

Taxonomic identification of *Diaporthe* species based on microsatellites

The eight polymorphic microsatellites provided an opportunity to discriminate *Diaporthe* species. The cluster analysis derived from Bruvo's distance showed that strains grouped into three clusters representing *D. longicolla, D. ueckerae,* and *D. unshiuensis* (Figure 2). Furthermore, the microsatellite marker C11 ((TGCC)6) discriminated the three species with five species-specific alleles. Isolates of *D. longicolla* had two alleles with 6 or 10 repeats; isolates of *D. unshiuensis* had two alleles with 7 or 9 repeats; and isolates of *D. ueckerae* had one allele with 3 repeats (Figure 3).

Population genotypic diversity and structures of *Diaporthe* species within four locations in Arkansas

Among 90 *Diaporthe* isolates included in this analysis, 88 unique MLGs were identified. One MLG was represented in two individuals (PL30, PL207) of *D. longicolla* across two sites,

while only one MLG of *D. unshiuensis* was duplicated (PL120, PL163) at the same site (Table 2; Figure 7). Even though most MLG data (97.8%) of *Diaporthe* species were unique, the minimum spanning network separated MLGs into the three *Diaporthe* species based on their genetic distance (Figure 4), which confirmed the ability of these loci to discriminate species (Figure 2). Because the sample size was different between populations, diversity was assessed on rarefied sub-samples of the data. Genotypic diversity indices revealed that populations of *D. longicolla* and *D. unshiuensis* were genetically highly diverse compared to *D. ueckerae*, possibly due to its low individual representation (Table 2, Figure 5). Furthermore, diversity indices of *Diaporthe* populations (90 individuals) indicated high genotypic diversity at all sites in Arkansas (Table 2, Figure 5). Additionally, *Diaporthe* genotypes from these sites did not genetically cluster as different lineages depending on regions. Specifically, genotypes of *D. longicolla* and *D. unshiuensis*) are likely mixed and not genetically clustered based on site of origin according to the minimum spanning network (Figure 4). However, *D. ueckerae,* and *D. unshiuensis* were not collected at some sites. Furthermore, populations of these species at the four locations (a total of nine populations) showed high genetic diversity, except for populations of *D. ueckerae* (Table 2, Figure 5). However, genotypes among these populations did not cluster based on geographic origin (Figure 4).

Linkage disequilibrium and AMOVA of *Diaporthe* species

The standard index of association ($\bar{r}_d$) and index of association (*Ia*) revealed differences significantly greater than zero in the clone corrected data of the entire set of 90 individuals at the four sites, which indicated a significant deviation from random mating and thus the hypothesis of random mating could be rejected. The value of $\bar{r}_d$ was 0.1 (P = 0.01) and *Ia* was 0.74 (P =0.01) (Figure 6), which indicates that loci were passed from parent populations to progenies in a non-

independent way. Furthermore, the hypothesis of random mating of *D. longicolla* was also rejected because non-zero values of the standard index of association ($\bar{r}_d$) and index of association (*Ia*) were observed in the clone corrected data (Figure 6). However, indices of linkage disequilibrium for *D. unshiuensis* populations were consistent with random mating.

To infer gene flow and population differentiation, genetic variation within and between levels of *Diaporthe* population hierarchy (sites and species) were assessed with AMOVA. High percentages of variation were observed between *Diaporthe* species and within sites. However, a low percentage of variation was observed within species across sites. Also, values of Phi between species and within sites were high compared with values between sites within species (Table 3). Because high variation between species and a low value within species between sites was noted, gene flow could have happened within individuals of the same species, but was unlikely to have occurred between *Diaporthe* species.

Mating types and vegetative incompatibility of *Diaporthe* species

Mating type was determined by amplifying part of the a1 box in MAT1-1-1, yielding a PCR fragment of 141 bp, and a portion of the HMG domain in MAT1-2-1, resulting in a PCR product of 229 bp. All examined isolates of the three *Diaporthe* species possessed both mating-types genes (Table 4). The existence of both mating types genes in the same isolate suggests that these three species are homothallic.

Vegetative compatibility tests were performed between *D. longicolla*, *D. ueckerae*, and *D. unshiuensis* isolates (Table 5, Figure 8). Pigmented barrage reaction lines between expanding colonies along the zone of mycelial contact were observed between isolates of the three species, which suggested genetic isolation between these species. Compartmentalized hyphal segments, vacuolated brown hyphae, and empty cells were observed in the reaction line between isolates

belonging to different species. This pigmented zone was also observed mainly among the isolates of *D. unshiuensis* and *D. ueckerae* and some isolates of *D. longicolla* (Table 6; Figure 9). However, dark lines in the contact zone were not observed among most isolates of *D. longicolla*.

Discussion

Genotyping by sequencing of microsatellites via next-generation sequencing to study the population genetics of *Diaporthe* species produced a large number of alleles (55) from eight loci with an average of 6.9 alleles per locus. GBS has distinct advantages over conventional approaches to assess SSR markers. For example, in another study, GBS approaches identified 44% more alleles than capillary electrophoresis (Darby et al., 2016). Capillary electrophoresis, in turn, typically identifies more alleles and provides higher resolution than agarose gel electrophoresis (e.g., Gupta et al., 2010; Stewart et al., 2011). In addition to being unable to resolve homoplasy, gel or capillary electrophoresis analyses are potentially affected by indels (deletion or insertion) in upstream or downstream regions of SSRs sequences. For instance, two *D. ueckerae* isolates had deletions in the C11 marker locus that could have resulted in false estimation of length via electrophoresis (Figure 3). Furthermore, in addition to resolving length homoplasy and accurately assessing size, genotyping by sequencing can multiplex hundreds of individuals and many loci into a single sequencing reaction (Darby et al., 2016).

Microsatellites are not only useful to assess population structure within a fungal species but can also be utilized to discriminate fungal species. In the current study, one SSR (C11) resolved isolates of *D. longicolla*, *D. ueckerae*, and *D. unshiuensis* at the species level. Likewise, SSRs of *Phytophthora* successfully distinguished some species of this genus (del Castillo-Munera et al., 2013). Although SSRs herein differentiated populations of *D. longicolla*, *D. ueckerae*, and *D. unshiuensis*, inferring phylogenetic relationships among these species using the

genetic distance of SSRs did not match phylogenetic inferences from the multigene tree (Figure 2, Chapter 2). However, based on the results of the minimum spanning network, *D. longicolla* may have diverged from the two *Diaporthe* species and formed different rooted lineages.

In this study, *Diaporthe* genotypes did not genetically cluster based on geographic origin, which could reflect a significant level of pathogen transport within Arkansas. Since *D. longicolla* can be associated with asymptomatic soybean seeds (Li, 2011), human-mediated transference could be an important component of redistribution in Arkansas. Insects provide another possible means of distribution; for example, stink bugs are capable of transporting *D. longicolla* and cause a yield loss of 20% in association with this fungus (Jones, 2013). More research will be required to resolve fundamental questions about the dynamics of pathogen distribution among *Diaporthe* species in Arkansas and other regions.

High levels of genetic variability and genotypic diversity were detected within populations of *D. longicolla* and *D. unshiuensis.* This is consistent with observations that other endophytic *Diaporthe* species display high genetic variability despite being isolated from leaves of the same plant at the same time (dos Santos et al., 2016). The high genetic variability of these species could be driven by selection pressures, such as adaptation to their environment. High levels of fungal genetic variability can facilitate the emergence of new genotypes during epidemics, the ability to survive difficult conditions, the ability to engage the sexual reproduction and thus reap the benefits of sexual recombination, and other valuable phenotypic traits (Milgroom, 2015).

Understanding sources of genetic diversity within plant pathogenic fungi is necessary for effective disease management. Numerous processes can influence genetic diversity, including sexual recombination, selection, random genetic drift, mutation, and migration (Milgroom,

2015). Because selection and random genetic drift do not directly increase genotypic diversity, sexual recombination, mutation, and migration could be important sources of genetic variability among *Diaporthe* species. However, the results of linkage disequilibrium indicate that populations of *D. longicolla* may be clonal based on the standard index of association ($\bar{r}_d$) and index of association (*Ia*) (Figure 6). On the other hand, however, sexual recombination could be driving genetic diversity among *D. unshiuensis* populations based on findings of linkage disequilibrium. In this study, mechanisms potentially driving genetic diversity in *D. ueckerae* could not be inferred because of few isolates were obtained.

AMOVA results indicated that the highest level of genetic variation occurred between species and within sites. However, the genetic differentiation within species across the sites was the lowest. These results suggest that a key source of genotypic diversity between Arkansas sites could be migration of genotypes and/or gene flow among strains. Consistent with this idea, high levels of vegetative compatibility were observed among isolates of *D. longicolla*. This observation supports the possibility of gene flow among individuals of this species in Arkansas, which could be reflected in the lack of grouping among isolates of *D. longicolla* based on geographic origin. However, most isolates of *D. unshiuensis* displayed vegetative incompatibility. Therefore, sexual recombination could be the key factor increasing genotypic diversity of this species. Vegetative incompatibility generally leads to programmed cell death after anastomosis, thus restricting parasexuality, which is a potential source of genetic variation in asexual fungi. However, parasexuality can occur between genetically incompatible isolates during certain conditions that might suppress vegetative incompatibility (Paoletti, 2016).

In conclusion, genotyping by sequencing of microsatellites successfully defined population compositions of endophytic *Diaporthe* isolates associated with soybean and

discriminated isolates of *D. longicolla, D. ueckerae,* and *D. unshiuensis* as clusters via genetic distance and species-specific alleles. Overall, high levels of genotypic diversity were observed among *D. longicolla and D. unshiuensis*, while diversity could not be fully assessed in *D. ueckerae* due to the low number of individuals sampled. Isolates of these species did not appear to group based on geographic origin, which indicates pathogen transport potential contributes to genetic diversity. These findings expand the current knowledge about genetic variation and population dynamics among *Diaporthe* species associated with soybean in Arkansas and can potentially contribute to the development of disease management strategies.

Literature Cited

Allen, T. W., Bissonnette, K., Bradley, C. A., Damicone, J. P., Dufault, N. S., Faske, T. R., . . . Young, H. (2019). *Southern United States soybean disease loss estimates for 2018.* Paper presented at the The 46 th Annual Meeting of the Southern Soybean Disease Workers, March 6 –7, Pensacola Beach, Florida.

Brown, A. H. D., Feldman, M. W., & Nevo, E. (1980). Multilocus structure of natural populations of *Hordeum spotaneum*. *Genetics,* 96(2), 523-536.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems,*1695(5), 1-9.

Darby, B. J., Erickson, S. F., Hervey, S. D., & Ellis-Felege, S. N. (2016). Digital fragment analysis of short tandem repeats by high-throughput amplicon sequencing. *Ecology and Evolution*, 6(13), 4502-4512. doi:10.1002/ece3.2221

del Castillo-Munera, J., Cardenas, M., Pinzon, A., Castaneda, A., Bernal, A. J., & Restrepo, S. (2013). Developing a taxonomic identification system of *Phytophthora* species based on microsatellites. *Revista Iberoamericana De Micologia,* 30(2), 88-95. doi:10.1016/j.riam.2012.11.002

Delmotte, F., Leterme, N., & Simon, J. C. (2001). Microsatellite allele sizing: Difference between automated capillary electrophoresis and manual technique. *Biotechniques,* 31(4), 810-814.

dos Santos, T. T., Leite, T. D., de Queiroz, C. B., de Araujo, E. F., Pereira, O. L., & de Queiroz, M. V. (2016). High genetic variability in endophytic fungi from the genus *Diaporthe* isolated from common bean (*Phaseolus vulgaris* L.) in Brazil. *Journal of Applied Microbiology,* 120(2), 388-401. doi:10.1111/jam.12985

Grunwald, N. J., Everhart, S. E., Knaus, B. J., & Kamvar, Z. N. (2017). Best Practices for Population Genetic Analyses. *Phytopathology,* 107(9), 1000-1010. doi:10.1094/phyto-12-16-0425-rvw

Grunwald, N. J., Goodwin, S. B., Milgroom, M. G., & Fry, W. E. (2003). Analysis of genotypic diversity data for populations of microorganisms. *Phytopathology,* 93(6), 738-746. doi:10.1094/phyto.2003.93.6.738

Guillin, E. A., Grijalba, P. E., de Oliveira, L. O., & Gottlieb, A. M. (2014). Specific boundaries between the causal agents of the soybean stem canker. *Tropical Plant Pathology*, 39(4), 316-325. doi:10.1590/s1982-56762014000400006

Gupta, V., Dorsey, G., Hubbard, A. E., Rosenthal, P. J., & Greenhouse, B. (2010). Gel versus capillary electrophoresis genotyping for categorizing treatment outcomes in two anti-malarial trials in Uganda. *Malaria Journal,* 9, 19-19. doi:10.1186/1475-2875-9-19

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., Carvalho, G. R. (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources,* 11, 123-136.

Jones, J. L. (2013). *Association between stink bug damage and the incidence of Phomopsis longicolla in Mississippi soybean production*: M.Sc. Thesis. Mississippi State University.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *Peer J.,* 2, e281.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., . . . Duran, C. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics,* 28(12), 1647-1649.

Kuhner, M. K., Beerli, P., Yamato, J., & Felsenstein, J. (2000). Usefulness of single nucleotide polymorphism data for estimating population parameters. *Genetics,* 156(1), 439-447.

Leboldus, J. M., Kinzer, K., Richards, J., Ya, Z., Yan, C., Friesen, T. L., & Brueggeman, R. (2015). Genotype-by-sequencing of the plant-pathogenic fungi *Pyrenophora teres* and *Sphaerulina musiva* utilizing Ion Torrent sequence technology. *Molecular Plant Pathology,* 16(6), 623-632.

Leslie, J. F., & Summerell, B. A. (2006). *The fusarium laboratory manual*. Ames, Iowa: Blackwell Pub.

Li, S. (2011). Phomopsis seed decay of soybean. In Sudaric A, (ed.), *Soybean-Molecular Aspects of Breeding*. Rijeka, Croatia: IntechOpen, 277–92.

Li, S. X., Darwish, O., Alkharouf, N., Matthews, B., Ji, P. S., Domier, L. L., . . . Bluhm, B. H. (2015). Draft genome sequence of *Phomopsis longicolla* isolate MSPL 10-6. *Genomics Data*, 3, 55-56. doi:10.1016/j.gdata.2014.11.007

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMB Net. Journal,* 17(1), 10-12.

Martins, W. S., Lucas, D. C. S., de Souza Neves, K. F., & Bertioli, D. J. (2009). WebSat-A web software for microsatellite marker development. *Bioinformation*, 3(6), 282.

Milgroom, M. G. (2015). *Population biology of plant pathogens: Genetics, ecology, and evolution*. St Paul, MN, USA: APS Press.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics,* 89(3), 583-590.

Nielsen, R. (2000). Estimation of population parameters and recombination rates from single nucleotide polymorphisms. *Genetics,* 154(2), 931-942.

Paoletti, M. (2016). Vegetative incompatibility in fungi: From recognition to cell death, whatever does the trick. *Fungal Biology Reviews,* 30(4), 152-162.

Pasqualotto, A. C., Denning, D. W., & Anderson, M. J. (2007). A cautionary tale: Lack of consistency in allele sizes between two laboratories for a published multilocus microsatellite typing system. *Journal of Clinical Microbiology,* 45(2), 522-528.

R Core Team. (2019). R: A Language and Environment for Statistical Computing. dim (ca533), 1(1358), 34.

Santos, J. M., Correia, V. G., & Phillips, A. J. (2010). Primers for mating-type diagnosis in *Diaporthe* and *Phomopsis*: Their use in teleomorph induction in vitro and biological species definition. *Fungal. Biol.,* 114(2-3), 255-270. doi:10.1016/j.funbio.2010.01.007

Schliep, K. P. (2011). phangorn: phylogenetic analysis in R. *Bioinformatics,* 27(4), 592-593. doi:10.1093/bioinformatics/btq706

Shannon, C. E., & Weaver, W. (1998). *The mathematical theory of communication*. Urbana, Illinois: University of Illinois press.

Simpson, E. H. (1949). Measurement of diversity. *Nature,*163(4148), 688-688.

Smith, J. M., Smith, N. H., O'Rourke, M., & Spratt, B. G. (1993). How clonal are bacteria? *Proceedings of the National Academy of Sciences,* 90(10), 4384-4388.

Stewart, S., Wickramasinghe, D., Dorrance, A. E., & Robertson, A. E. (2011). Comparison of three microsatellite analysis methods for detecting genetic diversity in *Phytophthora sojae* (Stramenopila: Oomycete). *Biotechnology Letters,* 33(11), 2217.

Stoddart, J. A., & Taylor, J. F. (1988). Genotypic diversity - estimation and prediction in samples. *Genetics,* 118(4), 705-711.

Suyama, Y., & Matsuki, Y. (2015). MIG-seq: an effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. *Scientific Reports*, 5, 12. doi:10.1038/srep16963

Vartia, S., Villanueva-Cañas, J. L., Finarelli, J., Farrell, E. D., Collins, P. C., Hughes, G. M., . . . Cross, T. F. (2016). A novel method of microsatellite genotyping-by-sequencing using individual combinatorial barcoding. *Royal Society Open Science*, 3(1), 150565.

Xu, J. (2006). Fundamentals of fungal molecular population genetic analyses. *Current Issues in Molecular Biology*, 8, 75-89.

Figures and Tables

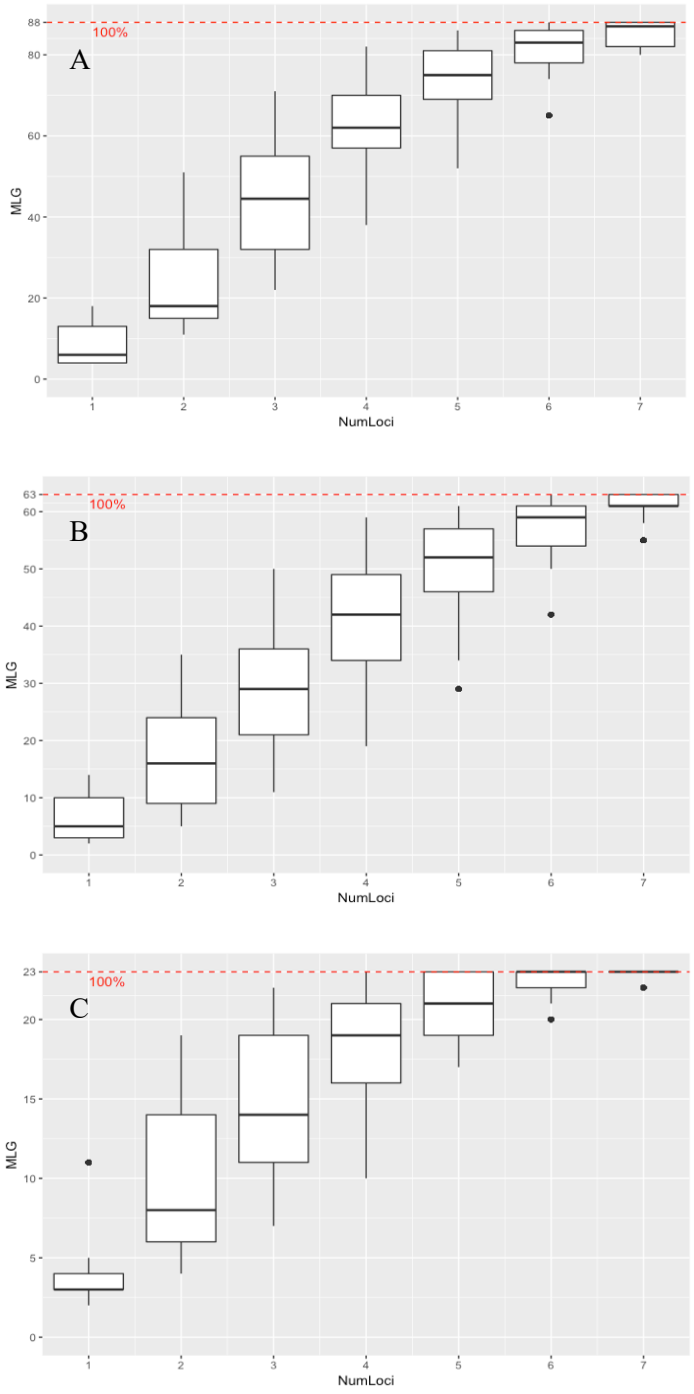


Figure 1. Genotype accumulation curve for populations of *Diaporthe* species in Arkansas. **(A)** Overall population of *Diaporthe* species. **(B)** Population of *D. longicolla.* **(C)** Population of *D. unshiuensis.*

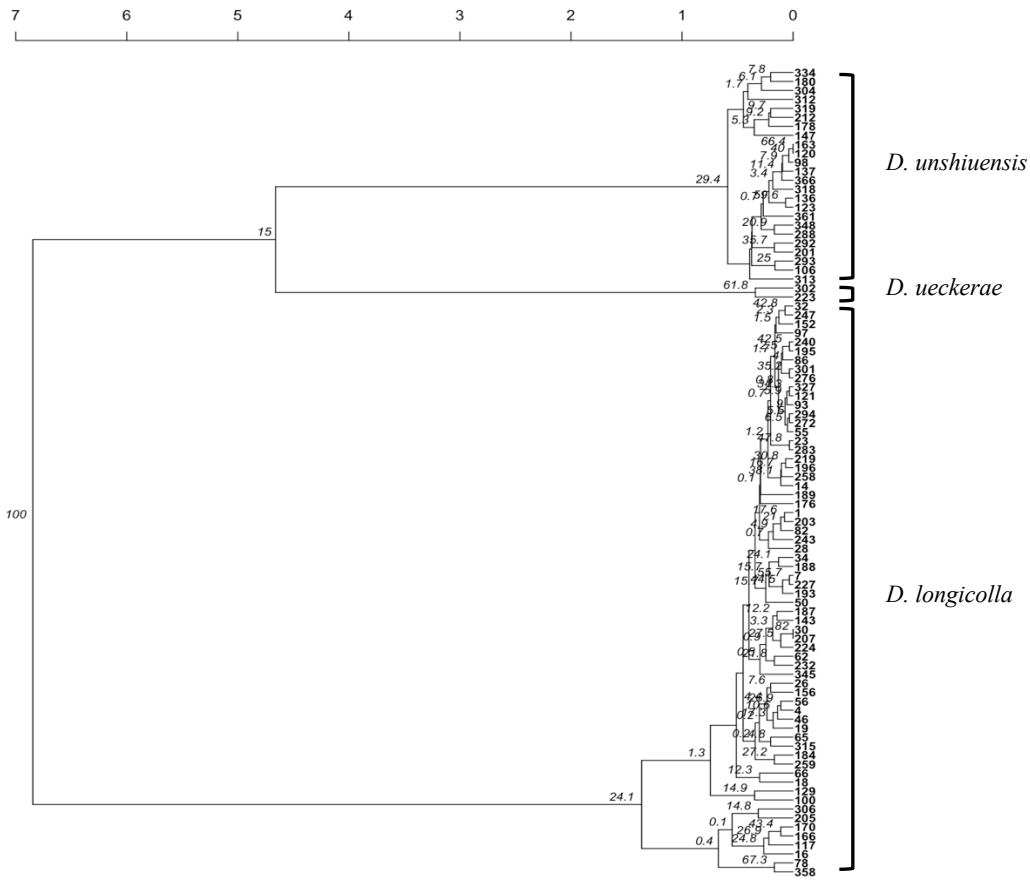


Figure 2. The dendrogram of cluster analysis based on genetic distance of the eight SSRs of *D. ueckerae*, *D. longicolla*, and *D. unshiuensis*.

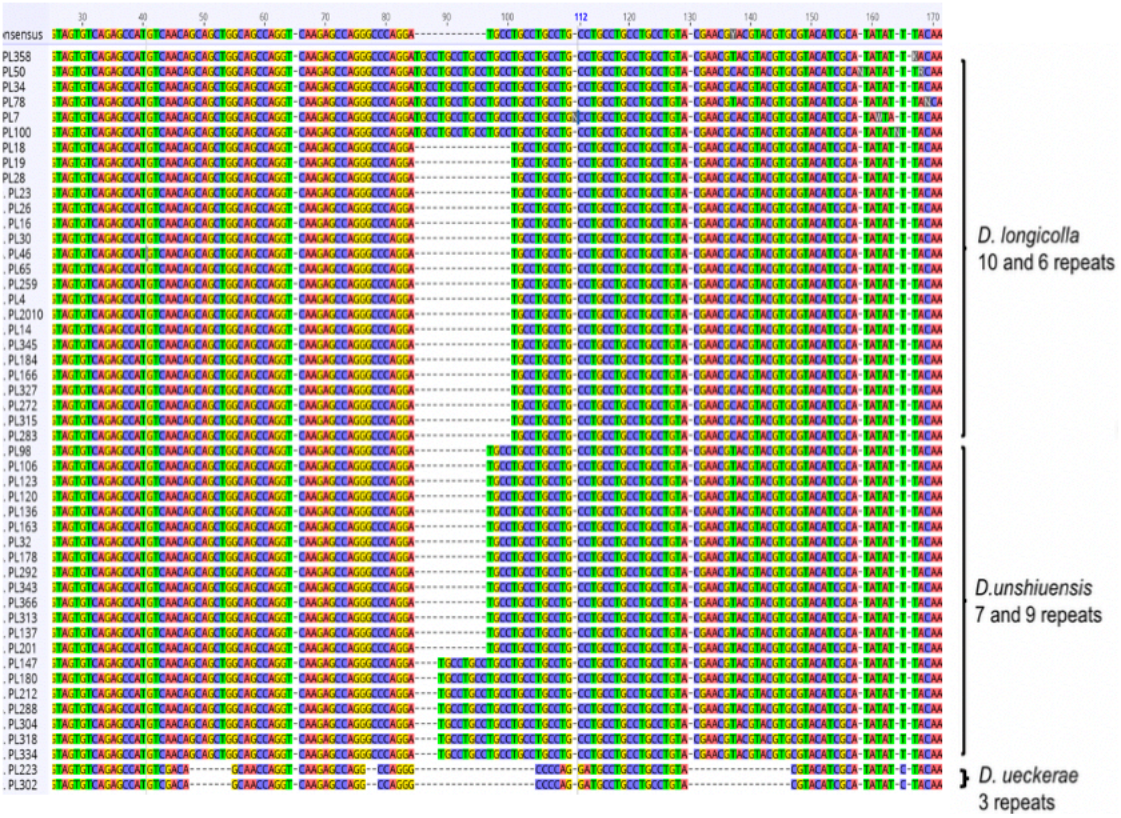


Figure 3. Sequences of the microsatellite C11((TGCC) 6) from selected isolates of *Diaporthe* species. Isolates of *D. longicolla* have two alleles, 6 and 10 repeats; isolates of *D. unshiuensis* have two alleles, 7 and 9 repeats; isolates of *D. ueckerae* have one allele, 3 repeats.

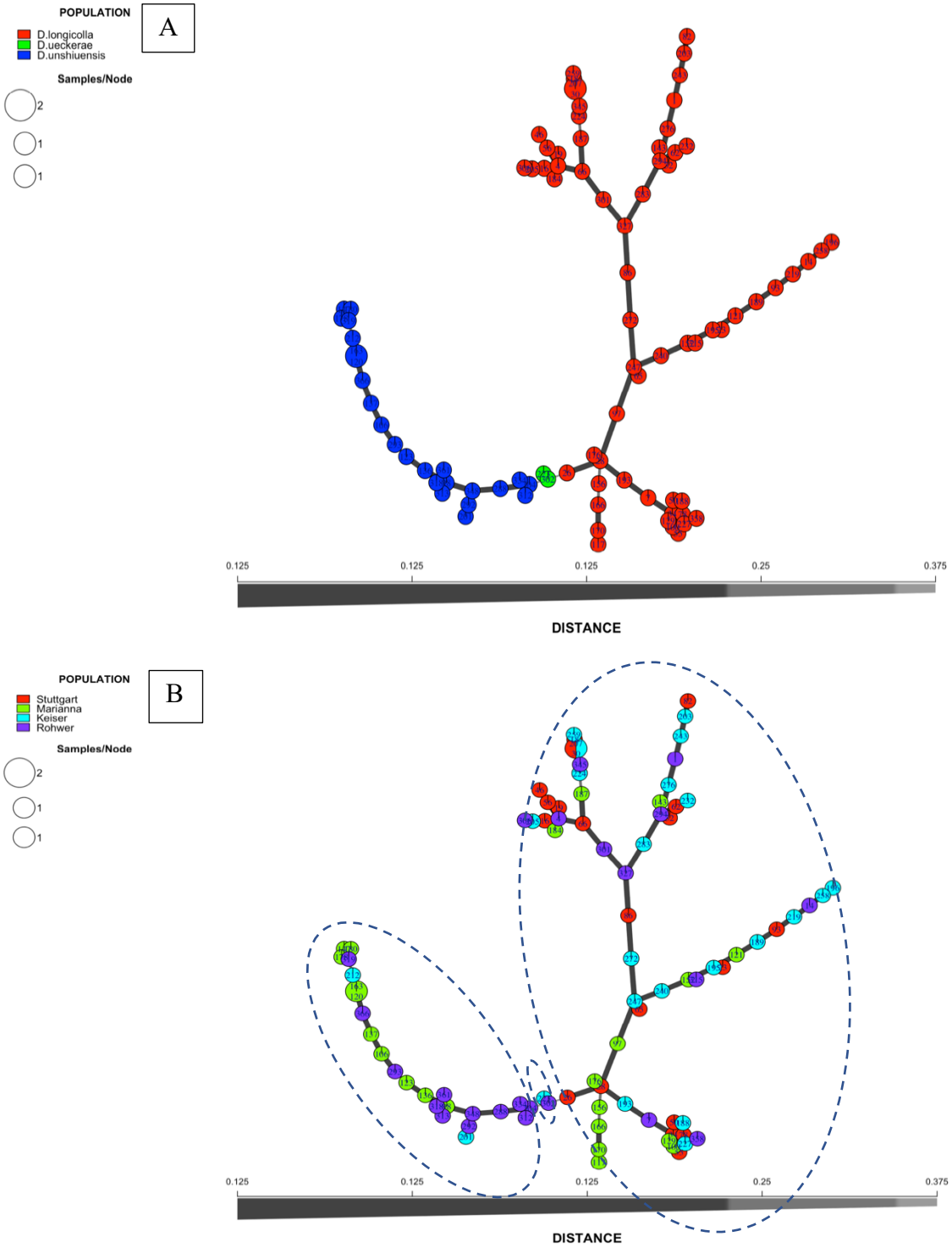


Figure 4. Minimum spanning network (MSN) of populations of *Diaporthe* isolates describing the relationships between multilocus genotypes after clone correction. **(A)** MSN for *Diaporthe* species. **(B)** MSN for *Diaporthe* isolates from four Arkansas locations.

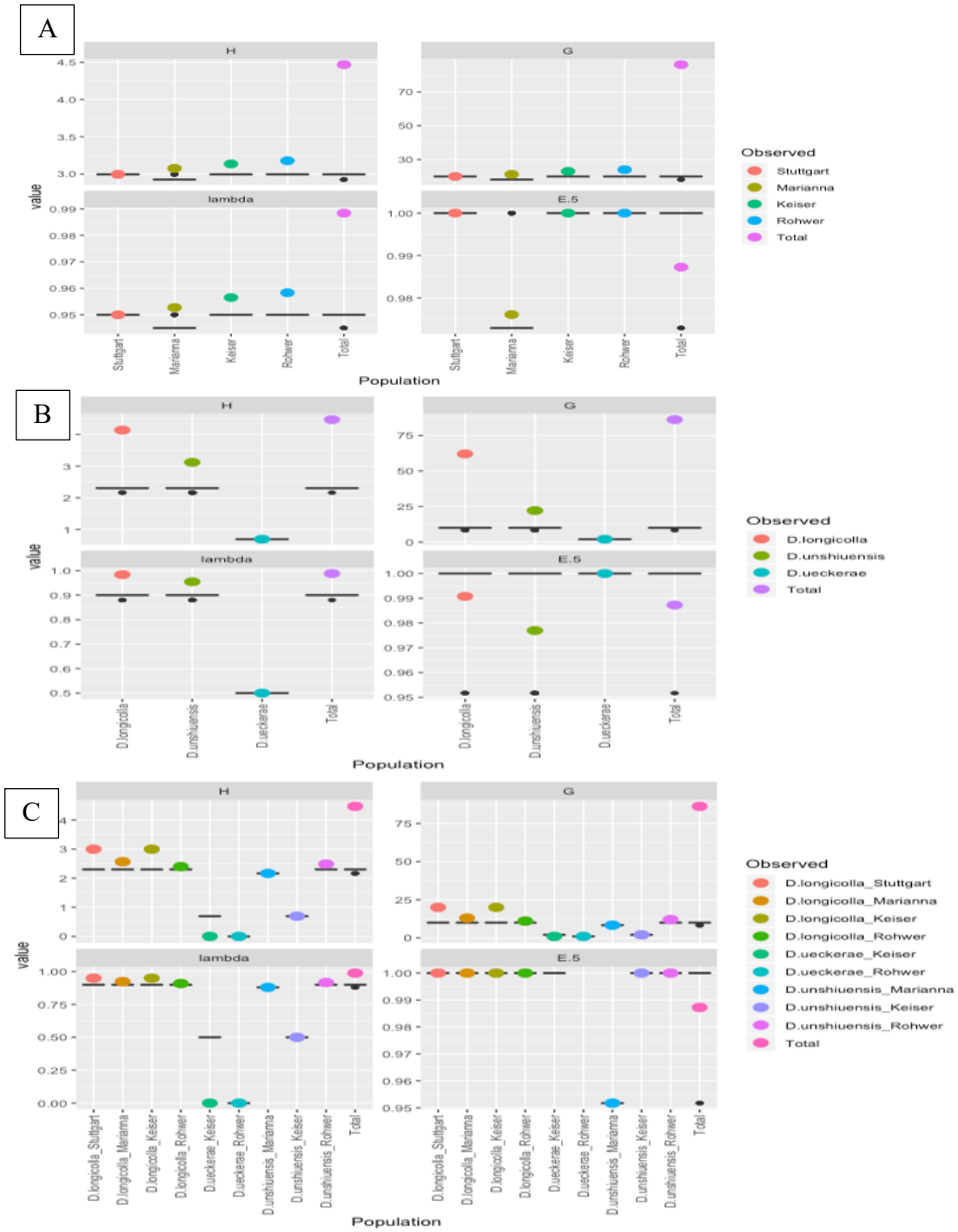


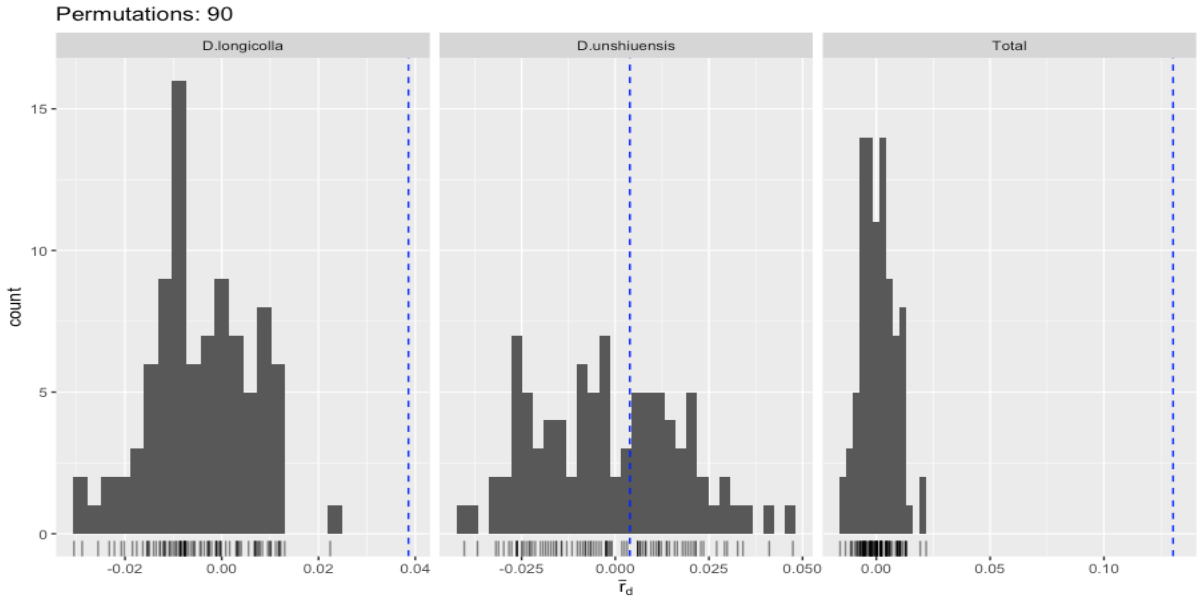
Figure 5. Comparison of genotypic diversity of *Diaporthe* species between: **(A)** Sampled sites, **(B)** *Diaporthe* species. **(C)** *Diaporthe* species - sampled site. The colored dots represent the observed statistics. The black box represents the values of rarefied samples estimated via bootstrapping with 1000 replicates. Samples for rarefaction: 20 for **(A)** and 10 each for **(B)** and **(C).**

H = Shannon-Wiener Index of MLG diversity

G = Stoddart and Taylor's Index of MLG diversity

Lambda = Simpson's Index

E.5 = Evenness



Pop	*D. longicolla*	*D. unshiuensis*	Total
Ia	0.269	0.027	0.906
p.Ia	0.011	0.407	0.011
rbarD	0.0386	0.0039	0.1304
p.rD	0.011	0.407	0.011

Figure. 6. Linkage disequilibrium of populations of *Diaporthe* species after clone correction.
*Values for *D. ueckerae* could not be plotted.
Ia = The index of association
p.Ia = P value of Ia
rbarD = The standardized index of association
p.rD = P value of rbarD

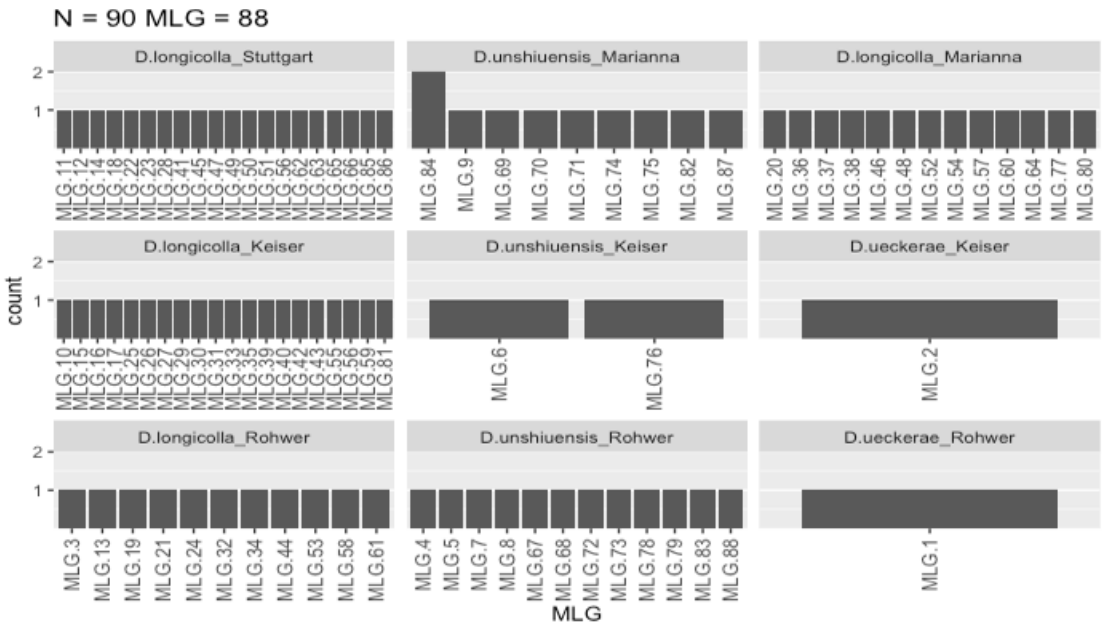


Figure 7. Multilocus genotype histogram (genotyped based multigene) for distribution of multilocus genotypes of *Diaporthe* species in Arkansas sites.

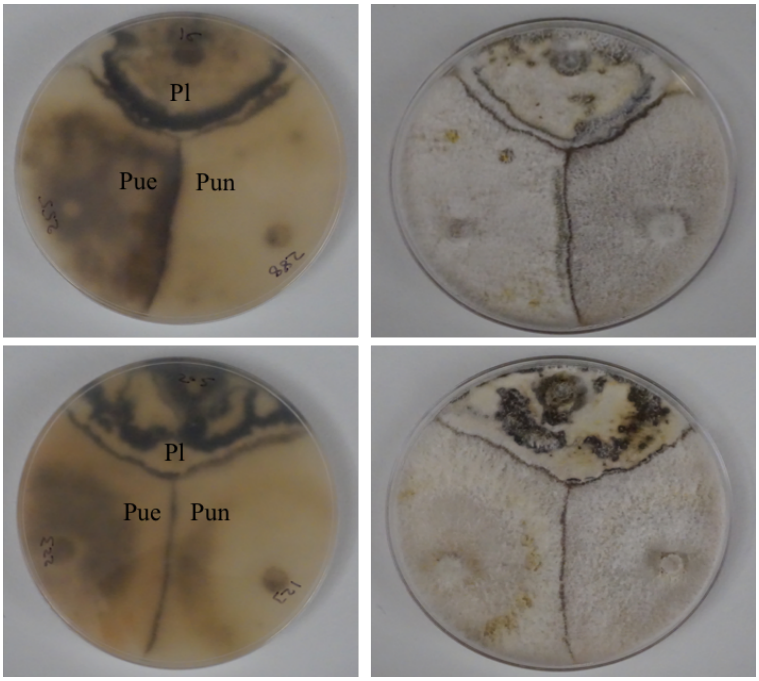


Figure 8. Vegetative incompatibility assay between *Diaporthe* species on PDA after two weeks. Two isolates per species; *D. longicolla* isolates (Pl), PL16 and PL205, *D. unshiuensis* isolates (Pun), PL288 and PL132, and *D. ueckerae* isolates (Pue), PL255 and 223.

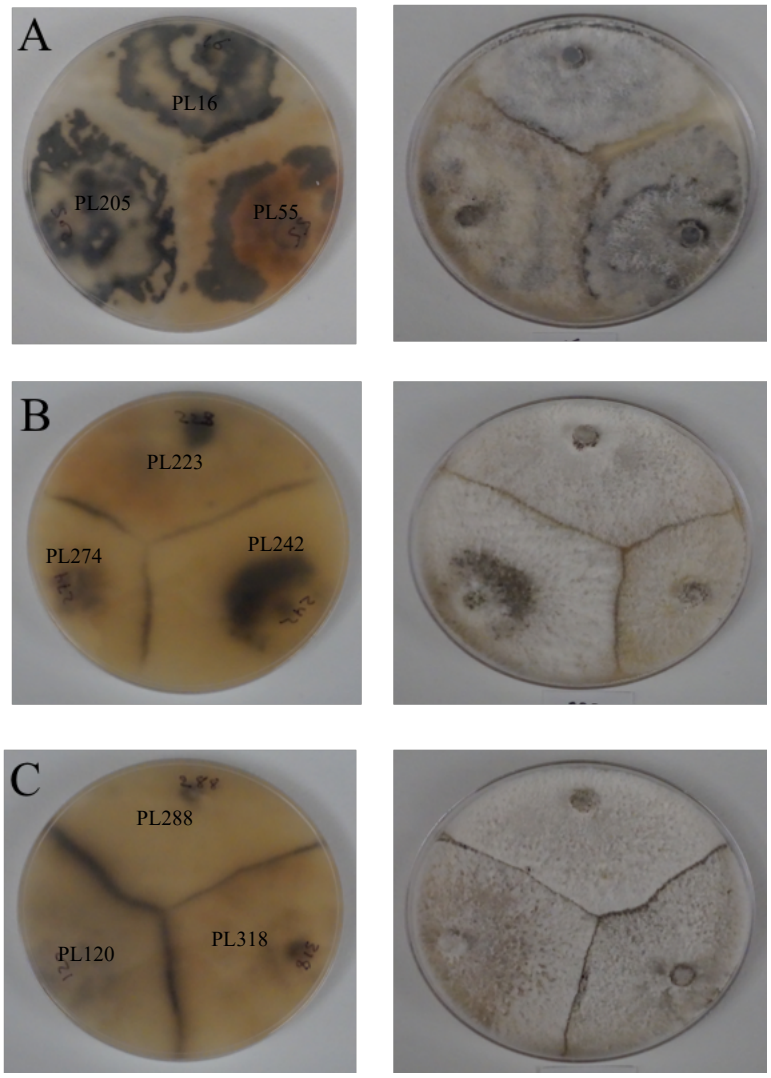


Figure 9. Vegetative incompatibility assays for *Diaporthe* species on PDA after two weeks growth, three isolates per species. **(A)** *D. longicolla* isolates PL16, PL55 and PL205. **(B)** *D. ueckerae* isolates PL242, PL274 and PL223. **(C)** *D. unshiuensis* isolates, PL288, PL318 and PL120.

Table 1. The genetic diversity of polymorphic SSR markers.

Locus	SSR	A	1-D	Hexp	E
A3	(CTGGCT)8	6.00	0.74	0.75	0.82
A4	(AGTC)6	3.00	0.52	0.53	0.80
A7	(TG)23	17.00	0.87	0.88	0.66
B10	(CA)7	4.00	0.65	0.66	0.84
C2	(AC)18	12.00	0.86	0.88	0.81
C11	(TGCC)6	5.00	0.59	0.60	0.65
D1	(AGTG)7	5.00	0.60	0.61	0.70
D12	(GTCT)6	3.00	0.60	0.61	0.86
Mean		6.88	0.68	0.69	0.77

A = Number of observed alleles
1-D = Simpson index
Hexp = Nei's 1978 gene diversity
E = Evenness

Table 2. Diversity statistics of clone-corrected populations of *Diaporthe* for **(A)** sampled sites **(B)** *Diaporthe* species **(C)** *Diaporthe* species within four locations.

A Pop	N	MLG	eMLG	SE	H	G	lambda	E.5
Stuttgart	20	20	20	0.00e+00	3.00	20	0.950	1.0
Marianna	22	22	20	7.45e-08	3.09	22	0.955	1.0
Keiser	23	23	20	0.00e+00	3.14	23	0.957	1.0
Rohwer	24	24	20	8.69e-08	3.18	24	0.958	1.0
Total	89	88	20	2.15e-01	4.47	87	0.989	0.993

B Pop	N	MLG	eMLG	SE	H	G	lambda	E.5
D. longicolla	63	63	10	1.38e-06	4.143	63	0.984	1.0
D. unshiuensis	23	23	10	5.03e-07	3.135	23	0.957	1.0
D. ueckerae	2	2	20	.00e+00	0.693	2	0.500	1.0
Total	88	88	10	0.00e+00	4.477	88	0.989	1.0

C Pops	N	MLG	eMLG	SE	H	G	lambda	E.5
*D. longicolla*_Stuttgart	20	20	10.00	0.00e+00	2.996	20.0	0.950	1.0
*D. unshiuensis*_Marianna	9	9	9.00	0.00e+00	2.197	9.0	0.889	1.0
*D. longicolla*_Marianna	13	13	10.00	7.30e-08	2.565	13.0	0.923	1.0
*D. longicolla*_Keiser	20	20	10.00	0.00e+00	2.996	20.0	0.950	1.0
*D. unshiuensis*_Keiser	2	2	2.00	0.00e+00	0.693	2.00	0.500	1.0
*D. ueckerae*_Keiser	1	1	1.00	0.00e+00	0.000	1.00	0.000	NaN
*D. longicolla*_Rohwer	11	11	10.00	0.00e+00	2.398	11.0	0.909	1.0
*D. unshiuensis*_Rohwer	12	12	10.00	0.00e+00	2.485	12.0	0.917	1.0
*D. ueckerae*_Rohwer	1	1	1.00	0.00e+00	0.000	1.00	0.000	NaN
Total	89	88	9.99	1.07e-01	4.473	87	0.989	0.993

N = Number of individuals observed
MLG = multilocus genotypes
eMLG = The number of expected MLG at the smallest sample size ≥ 10 based on rarefaction
SE = standard error from rarefaction
H = Shannon-Wiener Index of MLG diversity (Shannon, 2001)
G = Stoddart and Taylor's Index of MLG diversity (Stoddart & Taylor, 1988)
lambda = Simpson's Index (Simpson, 1949)
Hexp = Nei's unbiased gene diversity (Nei, 1978).
E.5 = Evenness, E5 (Pielou, 1975; Ludwig & Reynolds, 1988; Grünwald et al., 2003)
Ia = The index of association, IA (Brown, Feldman & Nevo, 1980; Smith et al., 1993)
rbarD = The standardized index of association, $\bar{r}_d$ [@].
NaN = could not be calculated

Table 3. Analysis of MOlecular VAriance (AMOVA) for simple sequence repeat data of *Diaporthe* species populations from four locations in Arkansas.

Hierarchical level	Sigma	Variation (%)	Phi(Φ)
Between species	3.38	42.66	0.426*
Between sites within species	0.50	0.62	0.0108
Within sites	4.49	56.72	0.433**
Total variations	7.91	100.00	

For this analysis, the data set was arranged into two hierarchical levels:
species and sites sampled.
Phi$_{st}$ (**Φ**) calculated for different hierarchical levels.
** = $p<0.01$, * = $p<0.0.5$

Table 4. Amplification of two mating-type genes of *Diaporthe* species.

Species	Strains	MAT1-1-1	MAT1-2-1
D. longicolla	PL16	+	+
D. longicolla	PL55	+	+
D. longicolla	PL78	+	+
D. longicolla	PL205	+	+
D. longicolla	PL243	+	+
D. unshiuensis	PL120	+	+
D. unshiuensis	PL123	+	+
D. unshiuensis	PL212	+	+
D. unshiuensis	PL288	+	+
D. unshiuensis	PL318	+	+
D. ueckerae	PL223	+	+
D. ueckerae	PL242	+	+
D. ueckerae	PL255	+	+
D. ueckerae	PL274	+	+
D. ueckerae	PL302	+	+

Table 5. Vegetative incompatibility between *Diaporthe* species.

	D. longicolla				
	Isolates	PL16	PL55	PL78	PL205	PL243
D. unshiuensis	PL120	+	+	+	+	+
	PL123	+	+	+	+	+
	PL212	+	+	+	+	+
	PL288	+	+	+	+	+
	PL318	+	+	+	+	+

	D. longicolla				
	Isolates	PL16	PL55	PL78	PL205	PL243
D. ueckerae	PL223	+	+	+	+	+
	PL242	-	+	+	+	+
	PL255	+	+	+	+	+
	PL274	+	+	+	+	+
	PL302	+	+	+	+	+

	D. ueckerae					
	Isolates	PL223	PL242	PL255	PL274	PL302
D. unshiuensis	PL120	+	+	+	+	+
	PL123	+	+	+	+	+
	PL212	+	+	+	+	+
	PL288	+	+	+	+	+
	PL318	+	+	+	+	+

Degree of incompatibility of *Diaporthe* species
– = no brown line observed
+ = brown line observed with vacuolated brown hyphae, and empty cells

Table 6. Vegetative incompatibility between arbitrarily selected isolates of *Diaporthe*.

	D. longicolla				
	Isolates	PL16	PL55	PL78	PL205	PL243
D. longicolla	PL16		-	-	- /+	-
	PL55			- / +	-	-
	PL78				-	-
	PL205					-
	PL243					

	D. ueckerae					
	Isolates	PL223	PL242	PL255	PL274	PL302
D. ueckerae	PL223		+	+	+	+
	PL242			+	+	-
	PL255				+	+
	PL274					+
	PL302					

	D. unshiuensis					
	Isolates	PL120	PL123	PL212	PL288	PL318
D. unshiuensis	PL120		+	+	+	+
	PL123			+	+	+
	PL212				+	+
	PL288					+
	PL318					

Degree of incompatibility of *Diaporthe* species
– = no brown line observed
+ = brown line observed with vacuolated brown hyphae, and empty cells.
- / + = brown line observed without vacuolated brown hyphae, and empty cells.

Supplementary Figures and Tables

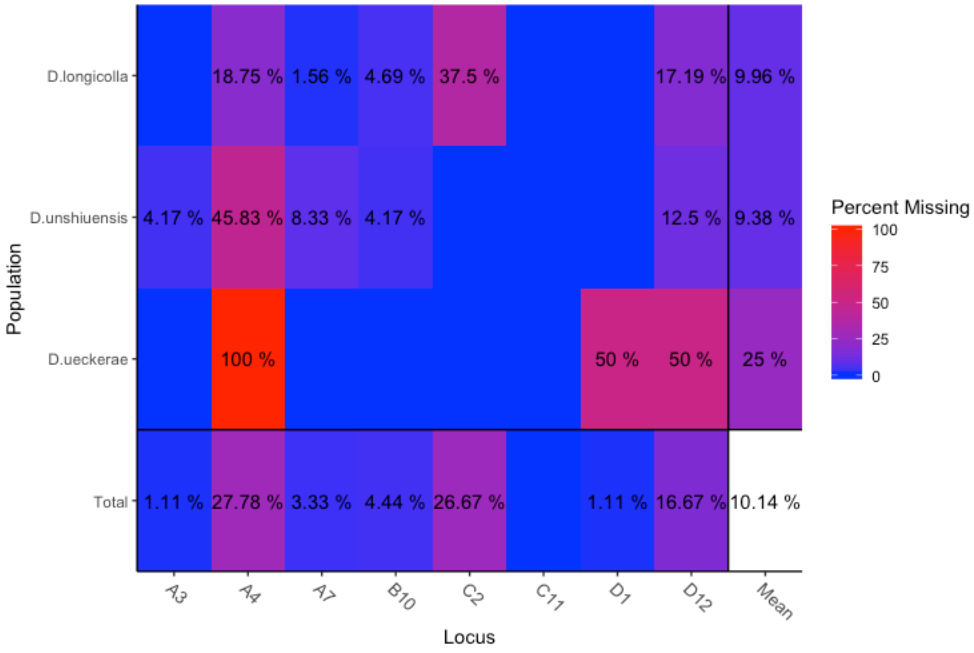


Figure 1S. The proportions of missing data per locus and population of *Diaporthe* in Arkansas.

Table 1S. *Diaporthe* isolates and their barcodes used in this chapter.

Number	Strain designation	Species	Site	Barcode
1	PL1	*D. longicolla*	Rohwer	TCTGCCTGTC
2	PL4	*D. longicolla*	Rohwer	CGATCGGTTC
3	PL7	*D. longicolla*	Rohwer	TCAGGAATAC
4	PL14	*D. longicolla*	Rohwer	CGAAGCGATTC
5	PL16	*D. longicolla*	Stuttgart	CTGCAAGTTC
6	PL18	*D. longicolla*	Stuttgart	CTAAGGTAAC
7	PL19	*D. longicolla*	Stuttgart	TAAGGAGAAC
8	PL23	*D. longicolla*	Stuttgart	TACCAAGATC
9	PL26	*D. longicolla*	Stuttgart	CAGAAGGAAC
10	PL28	*D. longicolla*	Stuttgart	AAGAGGATTC
11	PL30	*D. longicolla*	Stuttgart	TTCGTGATTC
12	PL32	*D. longicolla*	Stuttgart	CTGACCGAAC
13	PL34	*D. longicolla*	Stuttgart	TTCCGATAAC
14	PL46	*D. longicolla*	Stuttgart	TCCTCGAATC
15	PL50	*D. longicolla*	Stuttgart	TAGGTGGTTC
16	PL55	*D. longicolla*	Stuttgart	TCTAACGGAC
17	PL56	*D. longicolla*	Stuttgart	AACCTCATTC
18	PL62	*D. longicolla*	Stuttgart	TTCGAGACGC
19	PL65	*D. longicolla*	Stuttgart	TCTAGAGGTC
20	PL66	*D. longicolla*	Stuttgart	TCTGGATGAC
21	PL78	*D. longicolla*	Stuttgart	AGGCAATTGC
22	PL82	*D. longicolla*	Stuttgart	TTAGTCGGAC
23	PL86	*D. longicolla*	Stuttgart	CAGATCCATC
24	PL93	*D. longicolla*	Stuttgart	TGCCACGAAC
25	PL97	*D. longicolla*	Marianna	TTACAACCTC
26	PL98	*D. unshiuensis*	Marianna	CCTGAGATAC
27	PL100	*D. longicolla*	Marianna	AACCATCCGC
28	PL106	*D. unshiuensis*	Marianna	ATCCGGAATC
29	PL117	*D. longicolla*	Marianna	CGAGGTTATC
30	PL120	*D. unshiuensis*	Marianna	TTCTCATTGAAC
31	PL121	*D. longicolla*	Marianna	TCCAAGCTGC
32	PL123	*D. unshiuensis*	Marianna	TCTTACACAC
33	PL129	*D. longicolla*	Marianna	TCGCATCGTTC
34	PL136	*D. unshiuensis*	Marianna	TAAGCCATTGTC
35	PL137	*D. unshiuensis*	Marianna	AAGGAATCGTC
36	PL143	*D. longicolla*	Marianna	TGGAGGACGGAC
37	PL147	*D. unshiuensis*	Marianna	CTGACATAATC
38	PL152	*D. longicolla*	Marianna	AGCACGAATC

Table 1S. Cont.

Number	Isolates	Species	Site	Barcodes
39	PL156	*D. longicolla*	Marianna	TCAGTCCGAAC
40	PL163	*D. unshiuensis*	Marianna	TTCCACTTCGC
41	PL166	*D. longicolla*	Marianna	CTTGAGAATGTC
42	PL170	*D. longicolla*	Marianna	CTTGACACCGC
43	PL176	*D. longicolla*	Marianna	TTGGAGGCCAGC
44	PL178	*D. unshiuensis*	Marianna	TGGAGCTTCCTC
45	PL180	*D. unshiuensis*	Marianna	TAAGGCAACCAC
46	PL184	*D. longicolla*	Marianna	TAACAATCGGC
47	PL187	*D. longicolla*	Marianna	TTCTAAGAGAC
48	PL188	*D. longicolla*	Keiser	TCCTAACATAAC
49	PL189	*D. longicolla*	Keiser	CGGACAATGGC
50	PL193	*D. longicolla*	Keiser	TCCACCTCCTC
51	PL195	*D. longicolla*	Keiser	TCGACCACTC
52	PL196	*D. longicolla*	Keiser	TTGAGCCTATTC
53	PL201	*D. unshiuensis*	Keiser	CTGGCAATCCTC
54	PL203	*D. longicolla*	Keiser	TCTGGCAACGGC
55	PL205	*D. longicolla*	Keiser	TTCCTGCTTCAC
56	PL207	*D. longicolla*	Keiser	CCGGAGAATCGC
57	PL212	*D. unshiuensis*	Keiser	CAGCATTAATTC
58	PL219	*D. longicolla*	Keiser	TCCTAGAACAC
59	PL223	*D. ueckerae*	Keiser	TTCCTACCAGTC
60	PL224	*D. longicolla*	Keiser	TCCTTGATGTTC
61	PL227	*D. longicolla*	Keiser	TCTAGCTCTTC
62	PL232	*D. longicolla*	Keiser	TCACTCGGATC
63	PL240	*D. longicolla*	Keiser	CCTTAGAGTTC
64	PL243	*D. longicolla*	Keiser	CTGAGTTCCGAC
65	PL247	*D. longicolla*	Keiser	TCCTGGCACATC
66	PL258	*D. longicolla*	Keiser	TTCAATTGGC
67	PL259	*D. longicolla*	Keiser	CCGCAATCATC
68	PL272	*D. longicolla*	Keiser	TCAAGAAGTTC
69	PL276	*D. longicolla*	Keiser	CCTACTGGTC
70	PL283	*D. longicolla*	Keiser	TGAGGCTCCGAC
71	PL288	*D. unshiuensis*	Rohwer	CAGCCAATTCTC
72	PL292	*D. unshiuensis*	Rohwer	CGGAAGAACCTC
73	PL293	*D. unshiuensis*	Rohwer	CCTGGTTGTC
74	PL294	*D. longicolla*	Rohwer	TCGAAGGCAGGC
75	PL301	*D. longicolla*	Rohwer	CCTGCCATTCGC
76	PL302	*D. ueckerae*	Rohwer	CTTCCATAAC

Table 1S. Cont.

Number	Isolate	Species	Site	Barcode
77	PL304	*D. unshiuensis*	Rohwer	TTGGCATCTC
78	PL306	*D. longicolla*	Rohwer	CTAGGACATTC
79	PL312	*D. unshiuensis*	Rohwer	CCAGCCTCAAC
80	PL313	*D. unshiuensis*	Rohwer	CTTGGTTATTC
81	PL315	*D. longicolla*	Rohwer	TTGGCTGGAC
82	PL318	*D. unshiuensis*	Rohwer	CCGAACACTTC
83	PL319	*D. unshiuensis*	Rohwer	TCCTGAATCTC
84	PL327	*D. longicolla*	Rohwer	CTAACCACGGC
85	PL334	*D. unshiuensis*	Rohwer	CGGAAGGATGC
86	PL345	*D. longicolla*	Rohwer	CTAGGAACCGC
87	PL348	*D. unshiuensis*	Rohwer	CTTGTCCAATC
88	PL358	*D. longicolla*	Rohwer	TCCGACAAGC
89	PL361	*D. unshiuensis*	Rohwer	CGGACAGATC
90	PL366	*D. unshiuensis*	Rohwer	TTAAGCGGTC

[1] These barcodes (indexes) were used in the indexed primers to demultiplex and separate the sequences of *Diaporthe* isolates after pooling and sequencing the 10 SSRs of 90 isolates in one pooled library.

Table 2S. Microsatellites identified from the reference genome sequence of *D. longicolla* (strain PL2010).

Code	SSR[1]	Size[2]	Forward Primer	Reverse Primer
A3	(CTGGCT)8	219	CTAGTTTGATTACCCAGAAGCG	GAATCTCTTGGACAGAACTTGG
A4	(AGTC)6	390	GCAACCAAGACGACAAGACTAT	TCACTCTCACTCTCACTCTCCA
A7	(TG)23	305	TCTGCCGTACAAAAGGTACATA	TCTAGCAGGGTACAGGGATAAA
B3[3]	(AC)7	395	GGTCCGTATCTTGTGTAGAACC	TGTATTTCCCGCTATTTGAGAC
B10	(CA)7	278	TATCTATTGTTTACGGATGGGC	AGAACTAACTCACCGTTTCAGC
B11[3]	(GT)18	227	GGCTCTTTTACCCTTCCTACAC	GAACTCTCTCCTTGGGCTAGAT
C2	(AC)18	315	CACGGTTTGGCCTCTAGTATG	CTTGTCCAATAGTCATGCCACT
C11	(TGCC)6	288	AGAGGTAGTGTCAGAGCCATGT	GTCCATATTCGTTATAGCCGAG
D1	(AGTG)7	301	TGTGAGTGAATCTTAGCGAACT	CCTGCCTGTAAGGTACGAAG
D12	(GTCT)6	294	GCTCTCTGCGTATCCACACT	CCTGGTATTCCGTTATGTTTGA

[1] The number of repeats of SSRs in the reference genome of *D. longicolla*.
[2] The size of SSRs (bp) based on the reference genome of *D. longicolla*.
[3] Ten SSRs were sequenced and used for GBS in this study except B3 and B11, which were removed because of missing data.

Table 3S. Primers for the first round of PCR sequencing library construction.

Primer Name	Sequence
MIGA3F	CGC[1]TCT TCC GAT CTC TG[2]C TAG TTT GAT TAC CCA GAA GCG[3]
MIGB3F	CGC TCT TCC GAT CTC TG G GTC CGT ATC TTG TGT AGA ACC
MIGC11F	CGC TCT TCC GAT CTC TG A GAG GTA GTG TCA GAG CCA TGT
MIGA3R	TGC TCT TCC GAT CTG AC G AAT CTC TTG GAC AGA ACT TGG
MIGB3R	TGC TCT TCC GAT CTG AC T GTA TTT CCC GCT ATT TGA GAC
MIGC11R	TGC TCT TCC GAT CTG AC G TCC ATA TTC GTT ATA GCC GAG
MIGA4F	CGC TCT TCC GAT CTC TG G CAA CCA AGA CGA CAA GAC TAT
MIGB10F	CGC TCT TCC GAT CTC TG T ATC TAT TGT TTA CGG ATG GGC
MIGD1F	CGC TCT TCC GAT CTC TG T GTG AGT GAA TCT TAG CGA ACT
MIGA4R	TGC TCT TCC GAT CTG AC T CAC TCT CAC TCT CAC TCT CCA
MIGB10R	TGC TCT TCC GAT CTG AC A GAA CTA ACT CAC CGT TTC AGC
MIGD1R	TGC TCT TCC GAT CTG AC C CTG CCT GTA AGG TAC GAA G
MIGB11F	CGC TCT TCC GAT CTC TG G GCT CTT TTA CCC TTC CTA CAC
MIGB11R	TGC TCT TCC GAT CTG AC G AAC TCT CTC CTT GGG CTA GAT
MIGA7F	CGC TCT TCC GAT CTC TG T CTG CCG TAC AAA AGG TAC ATA
MIGC2F	CGC TCT TCC GAT CTC TG C ACG GTT TGG CCT CTA GTA TG
MIGD12F	CGC TCT TCC GAT CTC TG G CTC TCT GCG TAT CCA CAC T
MIGA7R	TGC TCT TCC GAT CTG AC T CTA GCA GGG TAC AGG GAT AAA
MIGC2R	TGC TCT TCC GAT CTG AC C TTG TCC AAT AGT CAT GCC ACT
MIGD12R	TGC TCT TCC GAT CTG AC C CTG GTA TTC CGT TAT GTT TGA

[1]Base pairs coded with orange color represented the anchor (3bp).
[2]Base pairs coded with red color represented primer tails (14bp).
[3]Base pairs coded with purple color represented SSR primers.

Table 4S. Reverse primer (a common primer without index) for the second round of PCR sequencing library construction.

Primer	Sequence
MIGR-trp	CCT CTC TAT GGG CAG TCG GTG AT[1]TGC[2]TCT TCC GAT CTG AC[3]

[1]Base pairs coded with black color represented the trp sequence.
[2]Base pairs coded with orange color represented the anchor (3bp).
[3]Base pairs coded with red color represented primer tails (14bp).

Table 5S. Primers for the second round of PCR sequencing library construction.

Forward Primers	Sequence
MIG-1	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG[1]CTAAGGTAACGAT[2]CGC[3]TCT TCC GAT CTC TG[4]
MIG-2	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TAAGGAGAACGAT CGC TCT TCC GAT CTC TG
MIG-3	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG AAGAGGATTCGAT CGC TCT TCC GAT CTC TG
MIG-4	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TACCAAGATCGAT CGC TCT TCC GAT CTC TG
MIG-5	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CAGAAGGAACGAT CGC TCT TCC GAT CTC TG
MIG-6	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTGCAAGTTCGAT CGC TCT TCC GAT CTC TG
MIG-7	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCGTGATTCGAT CGC TCT TCC GAT CTC TG
MIG-8	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCCGATAACGAT CGC TCT TCC GAT CTC TG
MIG-9	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TGAGCGGAACGAT CGC TCT TCC GAT CTC TG
MIG-10	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTGACCGAACGAT CGC TCT TCC GAT CTC TG
MIG-11	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCTCGAATCGAT CGC TCT TCC GAT CTC TG
MIG-12	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TAGGTGGTTCGAT CGC TCT TCC GAT CTC TG
MIG-13	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCTAACGGACGAT CGC TCT TCC GAT CTC TG
MIG-14	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTGGAGTGTCGAT CGC TCT TCC GAT CTC TG
MIG-15	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCTAGAGGTCGAT CGC TCT TCC GAT CTC TG
MIG-16	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCTGGATGACGAT CGC TCT TCC GAT CTC TG
MIG-17	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCTATTCGTCGAT CGC TCT TCC GAT CTC TG
MIG-18	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG AGGCAATTGCGAT CGC TCT TCC GAT CTC TG
MIG-19	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTAGTCGGACGAT CGC TCT TCC GAT CTC TG
MIG-20	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CAGATCCATCGAT CGC TCT TCC GAT CTC TG
MIG-21	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCGCAATTACGAT CGC TCT TCC GAT CTC TG
MIG-22	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCGAGACGCGAT CGC TCT TCC GAT CTC TG
MIG-23	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TGCCACGAACGAT CGC TCT TCC GAT CTC TG
MIG-24	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG AACCTCATTCGAT CGC TCT TCC GAT CTC TG
MIG-25	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCTGAGATACGAT CGC TCT TCC GAT CTC TG
MIG-26	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTACAACCTCGAT CGC TCT TCC GAT CTC TG
MIG-27	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG AACCATCCGCGAT CGC TCT TCC GAT CTC TG
MIG-28	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG ATCCGGAATCGAT CGC TCT TCC GAT CTC TG
MIG-29	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCGACCACTCGAT CGC TCT TCC GAT CTC TG
MIG-30	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CGAGGTTATCGAT CGC TCT TCC GAT CTC TG
MIG-31	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCAAGCTGCGAT CGC TCT TCC GAT CTC TG
MIG-32	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCTTACACACGAT CGC TCT TCC GAT CTC TG
MIG-33	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCTCATTGAACGAT CGC TCT TCC GAT CTC TG
MIG-34	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCGCATCGTTCGAT CGC TCT TCC GAT CTC TG
MIG-35	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TAAGCCATTGTCGAT CGC TCT TCC GAT CTC TG
MIG-36	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG AAGGAATCGTCGAT CGC TCT TCC GAT CTC TG
MIG-37	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTTGAGAATGTCGAT CGC TCT TCC GAT CTC TG

Table 5S. Cont.

Forward Primers	Sequence
MIG-38	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TGGAGGACGGACGAT CGC TCT TCC GAT CTC TG
MIG-39	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TAACAATCGGCGAT CGC TCT TCC GAT CTC TG
MIG-40	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTGACATAATCGAT CGC TCT TCC GAT CTC TG
MIG-41	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCCACTTCGCGAT CGC TCT TCC GAT CTC TG
MIG-42	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG AGCACGAATCGAT CGC TCT TCC GAT CTC TG
MIG-43	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTTGACACCGCGAT CGC TCT TCC GAT CTC TG
MIG-44	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTGGAGGCCAGCGAT CGC TCT TCC GAT CTC TG
MIG-45	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TGGAGCTTCCTCGAT CGC TCT TCC GAT CTC TG
MIG-46	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCAGTCCGAACGAT CGC TCT TCC GAT CTC TG
MIG-47	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TAAGGCAACCACGAT CGC TCT TCC GAT CTC TG
MIG-48	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCTAAGAGACGAT CGC TCT TCC GAT CTC TG
MIG-49	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCTAACATAACGAT CGC TCT TCC GAT CTC TG
MIG-50	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CGGACAATGGCGAT CGC TCT TCC GAT CTC TG
MIG-51	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTGAGCCTATTCGAT CGC TCT TCC GAT CTC TG
MIG-52	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCGCATGGAACGAT CGC TCT TCC GAT CTC TG
MIG-53	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTGGCAATCCTCGAT CGC TCT TCC GAT CTC TG
MIG-54	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCGGAGAATCGCGAT CGC TCT TCC GAT CTC TG
MIG-55	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCACCTCCTCGAT CGC TCT TCC GAT CTC TG
MIG-56	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CAGCATTAATTCGAT CGC TCT TCC GAT CTC TG
MIG-57	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCTGGCAACGGCGAT CGC TCT TCC GAT CTC TG
MIG-58	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCTAGAACACGAT CGC TCT TCC GAT CTC TG
MIG-59	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCTTGATGTTCGAT CGC TCT TCC GAT CTC TG
MIG-60	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCTAGCTCTTCGAT CGC TCT TCC GAT CTC TG
MIG-61	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCACTCGGATCGAT CGC TCT TCC GAT CTC TG
MIG-62	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCCTGCTTCACGAT CGC TCT TCC GAT CTC TG
MIG-63	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCTTAGAGTTCGAT CGC TCT TCC GAT CTC TG
MIG-64	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTGAGTTCCGACGAT CGC TCT TCC GAT CTC TG
MIG-65	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCTGGCACATCGAT CGC TCT TCC GAT CTC TG
MIG-66	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCGCAATCATCGAT CGC TCT TCC GAT CTC TG
MIG-67	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCCTACCAGTCGAT CGC TCT TCC GAT CTC TG
MIG-68	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCAAGAAGTTCGAT CGC TCT TCC GAT CTC TG
MIG-69	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCAATTGGCGAT CGC TCT TCC GAT CTC TG
MIG-70	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCTACTGGTCGAT CGC TCT TCC GAT CTC TG
MIG-71	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TGAGGCTCCGACGAT CGC TCT TCC GAT CTC TG
MIG-72	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CGAAGGCCACACGAT CGC TCT TCC GAT CTC TG
MIG-73	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCTGCCTGTCGAT CGC TCT TCC GAT CTC TG
MIG-74	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CGATCGGTTCGAT CGC TCT TCC GAT CTC TG

Table 5S. Cont.

Forward Primers	Sequence
MIG-75	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCAGGAATACGAT CGC TCT TCC GAT CTC TG
MIG-76	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CGGAAGAACCTCGAT CGC TCT TCC GAT CTC TG
MIG-77	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CGAAGCGATTCGAT CGC TCT TCC GAT CTC TG
MIG-78	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CAGCCAATTCTCGAT CGC TCT TCC GAT CTC TG
MIG-79	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCTGGTTGTCGAT CGC TCT TCC GAT CTC TG
MIG-80	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCGAAGGCAGGCGAT CGC TCT TCC GAT CTC TG
MIG-81	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCTGCCATTCGCGAT CGC TCT TCC GAT CTC TG
MIG-82	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTGGCATCTCGAT CGC TCT TCC GAT CTC TG
MIG-83	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTAGGACATTCGAT CGC TCT TCC GAT CTC TG
MIG-84	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTTCCATAACGAT CGC TCT TCC GAT CTC TG
MIG-85	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCAGCCTCAACGAT CGC TCT TCC GAT CTC TG
MIG-86	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTTGGTTATTCGAT CGC TCT TCC GAT CTC TG
MIG-87	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTGGCTGGACGAT CGC TCT TCC GAT CTC TG
MIG-88	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CCGAACACTTCGAT CGC TCT TCC GAT CTC TG
MIG-89	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCTGAATCTCGAT CGC TCT TCC GAT CTC TG
MIG-90	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTAACCACGGCGAT CGC TCT TCC GAT CTC TG
MIG-91	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CGGAAGGATGCGAT CGC TCT TCC GAT CTC TG
MIG-92	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTAGGAACCGCGAT CGC TCT TCC GAT CTC TG
MIG-93	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CTTGTCCAATCGAT CGC TCT TCC GAT CTC TG
MIG-94	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCGACAAGCGAT CGC TCT TCC GAT CTC TG
MIG-95	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG CGGACAGATCGAT CGC TCT TCC GAT CTC TG
MIG-96	CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTAAGCGGTCGAT CGC TCT TCC GAT CTC TG

[1]Base pairs coded with black color represented Adapter A (30 bp).
[2]Base pairs coded with blue color represented the barcode (10-12 bp).
[3]Base pairs coded with orange color represented the anchor (3 bp).
[4]Base pairs coded with red color represented primer tails (14 bp).

Table 6S. Mating type primers used in this study (Santos et al., 2010).

Primers	Sequences[1]
MAT1-1-1FW	5-GCA AMI GTK TIK ACT CAC A-3
MAT1-1-1RV	5-GTC TMT GAC CAR GAC CAT G-3
MAT1-2-1FW	5-GCC CKC CYAAYC CAT TCA TC-3
MAT1-2-1RV	5-TTG ACY TCA GAA GAC TTG CGT G-3

[1] Wobble bases are comprised of the following nucleotide combinations:
M = A/C
K = G/T
R = A/G
Y = C/T
I = Inosine

Chapter 4: Forward Genetic Screen for Pathogenicity Genes of the Fungus *Diaporthe longicolla* Causing Phomopsis Seed Decay of Soybean

Abstract

Phomopsis seed decay of soybean is an economically important disease in the U.S. The disease is predominantly caused by *Diaporthe longicolla*. Currently, the molecular basis of Phomopsis seed decay is poorly understood. The objective of this study was to identify genes of *D. longicolla* involved in the colonization of soybean seeds and stem necrosis. Random insertional mutagenesis via *Agrobacterium tumefaciens* mediated transformation (ATMT) generated 1,251 mutants of the pathogen. Two mutants with visually reduced seed colonization were selected for further study from a forward genetic screen. Target enrichment sequencing identified a single site of *Agrobacterium*-mediated T-DNA insertion in each mutant. In one mutant (PLM2739), a T-DNA insertion disrupted a gene encoding a putative serine threonine protein. In the other mutant (PLM1983), a T-DNA insertion was identified in the putative promoter region of a gene (g2420.t1) predicted to encode a cytochrome P450. Virulence assays indicated that both mutants were impaired in seed colonization and the ability to induce necrotic lesions on stems. Additionally, neither mutant produced pycnidia on four different culture media. Reintroducing g2420.t1 into strain PLM1983 via ATMT partially restored virulence and conidiation. Also, the relative expression of g2420.t1 was significantly lower in the mutant compared with the wild- type. However, genetic complementation of strain PLM2739 was not successful. These findings suggest a critical role for g2420.t1 in pathogenicity and asexual reproduction in *D. longicolla*.

Introduction

Diaporthe longicolla (Hobbs) J.M. Santos, Vrandečić & A.J.L. Phillips (syn. *Phomopsis longicolla* Hobbs) is a prevalent pathogen of soybean. This pathogen is considered the primary cause of Phomopsis seed decay (PSD), which is an economically important soybean disease in the U.S. Midsouth, including Arkansas (Li, 2011). In recent years, PSD caused substantial soybean yield losses in Arkansas (Allen et al., 2018; Allen et al., 2019). Symptoms include shriveled, elongated, or cracked seeds, often with a chalky-white appearance in addition to symptomless infection. As consequences of seed infection, seeds fail to germinate or germinate more slowly than healthy seeds (Kulik & Sinclair, 1999; Sinclair, 1993). Therefore, PSD is considered to be one of the most important factors affecting soybean seed germination and vigor (Gillen et al., 2012). *D. longicolla* can also potentially cause stem blight and canker (Cui et al., 2009; Tolbert & Spurlock, 2017) and leaf spot disease (Xue et al., 2015). In addition to causing disease, *D. longicolla* produces cytotoxic and antimicrobial secondary metabolites, such as phomoxanthones and dicerandrols (Isaka et al., 2001; Lin et al., 2010). Consequently, consuming infected seeds and derived products could potentially threaten human and animal health.

Management of PSD is problematic. Fungicides have been used to control the disease, but this approach is not always successful (TeKrony et al., 1985; Wrather et al., 2004). Agricultural practices such as conventional tillage and rotation with non-legume crops have also provided inconsistent control (Li et al., 2018). Another method used to control PSD is host resistance. Durable resistance to plant disease is highly cost effective (Agrios, 2005). Efforts have been made to identify soybean lines with resistance to PSD by screening soybean germplasm in the field (Li et al., 2015; Li et al., 2017). Resistance to PSD was conferred by a single dominant gene in some soybean lines, and another line was found to have two

complementary dominant genes associated with strong resistance to PSD (Jackson et al., 2009; Smith et al., 2008). However, widespread deployment of genetic resistance to PSD is not yet available in commercial cultivars.

Previous studies have addressed the mode of seed infection and the molecular basis of pathogenicity of *D. longicolla*. The fungus can form appressoria and penetrate the cell wall directly. It can infect soybean pods at any time by invading the ovule and developing seeds via the funiculus and hilum. Although *D. longicolla* has not been observed to infect immature soybean pods via natural openings, the fungus can access soybean seed coats through natural openings in mature seeds (Baker et al., 1987; Roy & Abney, 1988). Therefore, maximal seed infection by *D. longicolla* may occur after seeds become physiologically mature (Hepperly & Sinclair, 1980; Kmetz et al.,1978; McGee, 1984). However, genetic mechanisms underlying pathogenicity of *D. longicolla* have not been investigated extensively (Zaccaron, 2019) and could be somewhat complicated due to wide host range and different lifestyles. For example, *D. longicolla* can associate with weeds and crops as an endophyte or pathogen (Li, 2011; Mengistu et al., 2007). Therefore, more efforts are required to clearly understand pathogenesis in this organism.

Genetic techniques and new resources to facilitate studying pathogenesis in *D. longicolla* have recently become available. Forward genetic screening has been a powerful tool to determine genes and mutations that underlie phenotypes of interest and characterize functionally unknown genes (Schneeberger, 2014). Li et al. (2013) developed a robust technique to create random insertional transformants of *D. longicolla* that could be used with forward and reverse genetic screening. Furthermore, a reference genome of *D. longicolla* has been sequenced and annotated (Darwish et al., 2016a; Li et al., 2015). However, functional characterization of

pathogenicity genes is limited in *D. longicolla* (Zaccaron, 2019). In the current study, *Agrobacterium tumefaciens* mediated transformation (ATMT) and forward genetic screening were adopted to identify genes of *D. longicolla* underlying pathogenicity and asexual reproduction.

Materials and Methods

Strains and vectors

D. longicolla strain PL2010 originating from Crawford County, Arkansas, USA, (Li et al., 2013) was used as the wild-type isolate in this study. *Agrobacterium tumefaciens* strain AGL-1 (Lazo et al., 1991) was used to create transformants with the binary vector pBHt2_SGFP (derived from pBHt2) (Mullins et al., 2001), which contains screenable and selectable markers: the *sGFP* gene driven by the *ToxA* promoter (Lorang et al., 2001) and the hygromycin resistance gene (*hph*) driven by the *trpC* promoter (Staben et al., 1989) (Figure 1A). To complement mutants, pBYR48 was used as a backbone to construct vectors containing genes of interest (Figure 1B).

Creation of randomly tagged *D. longicolla* mutants and fungal transformation

A. tumefaciens mediated transformation (ATMT) was used to create randomly inserted transformants of wild-type strain PL2010 of *D. longicolla* according to the protocol of Li et al. (2013). Briefly, *A. tumefaciens* strain AGL-1 with vector pBHt2-sGFP was plated on luria broth (LB) agar medium with carbenicillin (50 µg/ml) and kanamycin (50 µg/ml) and incubated at 28 °C for 48 h. One colony of *A. tumefaciens* was picked to inoculate a 15-ml tube with 5 ml LB containing 50 µg/ml carbenicillin and 50 µg/ml kanamycin, which was incubated with shaking at 150 rpm and 28 °C for 48 h. Cells were diluted to OD_{600} of 0.15 in 5 ml of induction minimal medium (IMM) containing 200 µM acetosyringone. When the OD_{600} reached 0.25, 100 µl of *A.*

tumefaciens cell suspension was mixed with 100 µl of suspended conidia (10^6 conidia/ml) from strain PL2010 of *D. longicolla*. The mixture was spread immediately onto sterile cellophane membranes, overlaid on IMM agar plates (6 cm) containing 200 µM acetosyringone, and incubated for three days at 25 °C in darkness. Then, cellophane membranes were inverted and transferred to 0.2× PDA plates (9 cm) containing hygromycin B (100 µg/ml) and cefotaxime (200 µg/ml). After 48 h, the cellophane membranes were removed and discarded. Transformants formed visible colonies within two days. Colonies of mutants were picked and plated in 96-well plates containing 0.2× PDA and hygromycin B (100 µg/ml).

Screening of random insertional transformants of *D. longicolla*

The early-maturing soybean variety Traff (PI 470930) was selected to evaluate transformants of *D. longicolla* for seed colonization. Seeds were inoculated with 1,251 mutants of *D. longicolla* created as described above following the modified protocol of Xue et al. (2006). Briefly, mycelial suspensions for each mutant and the wild type strain were prepared by harvesting mycelium from 1 cm^2 of a colony (age 5 days) in 1 ml distilled water, and homogenizing tissue with a bead beater for 3 min in a 2-ml microcentrifuge tube. Mycelial suspensions (0.5 ml) were uniformly spread on 6-cm PDA plates and incubated for 24 h at 25 °C before inoculation of soybean seeds. To collect soybean seeds for inoculation, yellow pods from cultivar Traff were washed in tap water and sterilized by submersion in 70% ethanol for 30 s, 0.8% bleach for 1 min, and sterilized distilled water twice. Seeds were extracted from pods and sterilized by submersion in 70% ethanol for 30 s, 0.8% bleach for 1 min, and sterilized distilled water twice. Five seeds were placed on the surface of mycelia growing in 6-cm petri dishes containing PDA, inoculated as described above. Petri dishes were incubated in plastic bags with continuous lighting at 25 °C. After 5 days, visual assessments of seed colonization were rated

with a scale (0-10). Zero represented seeds with no observable fungal growth, and a value of ten indicated seeds were engulfed by fungal growth (Xue et al., 2006). Mutants of *D. longicolla* that scored 1-9 were selected for another seed colonization screening with three replicates. Mutants with impaired seed colonization were re-evaluated with a cut-stem pathogenicity assay following the protocol of Li et al. (2010). Briefly, 2-week-old soybean stems were cut below the first trifoliate node and inoculated with 5-mm mycelial discs taken from 10-day-old colonies on APDA with the large ends of 200-µl disposable micropipette tips. The negative control treatment was inoculated with micropipette tips containing plugs of uninoculated APDA. The experimental units were arranged in a complete randomized block design with six replicates (e.g., pots), with four seedlings per pot (10 x 10 x 9 cm). Micropipette tips were removed two days after inoculation and stem lesion lengths were assessed 7 days following the inoculation. The experiment was conducted twice. The growth and conidiation of five interesting mutants were evaluated on four different media to assess the effect of disrupted genes on these parameters.

DNA extraction, preparing libraries and sequencing for identifying T-DNA sites

Genomic DNA of mutants PLM1983 and PLM2739 were extracted with a DNeasy Plant Mini Kit according to manufacturer's instructions (QIAGEN, Valencia, California, USA). To prepare libraries, the NEBNext® Fast DNA Fragmentation and Library Prep Set for Ion Torrent PGM was used with an initial DNA amount of 500 ng for each library. Fragmentation and end repair of DNA, and preparation of adapter ligated DNA were performed following manufacturer's instructions (New England Biolabs, Ipswich, MA). Probe capture method for capturing 360 bp from the left and the right border of the target insertional T-DNA was done via a double hybridization method adapted from Schmitt et al. (2015) with biotinylated probes that were designed and ordered from IDT (Integrated DNA Technologies, Coralville, IA).

Briefly, 500 ng of each library was mixed with 1 µl of each of xGen Universal Blocking Oligo 1 (IT-P1) and xGen Universal Blocking Oligo 2 (IT-A) (Integrated DNA Technologies). After the mixture was lyophilized with a Savant Speed Vac Concentrator SVC100H (Thermo Fisher Scientific, Waltham, MA, USA), the following reagents were added to each lyophilized sample: 8.5 µl of xGen 2× Hybridization Buffer (Integrated DNA Technologies), 2.7 µl of xGen Hybridization Buffer Enhancer (Integrated DNA Technologies), and 1.8 µl nuclease-free water. Samples were incubated at 95 °C for 10 min. Then, after adding 3 pmol of biotinylated probes, reactions were incubated at 65 °C for 4 h. Next, 75 µl of Dynabeads M-270 Streptavidin beads (Thermo Fisher Scientific Inc., Waltham, MA) were added and the mixtures were incubated and washed according to the protocol described in the hybridization capture of DNA libraries using xGen Lockdown Probes and Reagents (Integrated DNA Technologies). The probed libraries were amplified for 14 cycles as described in the NEBNext® Fast DNA Fragmentation and Library Prep Set for Ion Torrent® (New England Biolabs, Ipswich, MA) and purified with 1× volume of Agencourt AMPure XP beads (Beckman Coulter, Brea, CA). An additional round of insertional T-DNA target enrichment was performed from the purified libraries with blocking oligos and biotinylated probes following the method described above. Subsequently, the product was cleaned up with Agencourt AMPure XP beads as described above. The two libraries were pooled in equal molar concentrations and sequenced on an Ion Torrent Personal Genome Machine (PGM) with an Ion 318 chip kit V2 (Thermo Fisher Scientific Inc., Waltham, MA).

Bioinformatics and identification of the insertional sites of T-DNA

The sequenced reads were mapped to plasmid pBHT2_sGFP with BWA-MEM version 0.7.12 (Li, 2013). The reads that successfully mapped to the plasmid were mapped to the *P. longicolla* MSPL 10-6 genome (Li et al., 2015) with BWA-MEM version 0.7.12. SAMtools

version 0.1.19 (Li et al., 2009) was used to filter unmapped reads and secondary alignments. Mapped reads were grouped with the *merge* subcommand within BED tools suite v2.26.0 (Quinlan & Hall, 2010). Overlapping and/or book-ended reads were grouped into the same group. Read mapping was visualized with IGV version 2.3.57 (Thorvaldsdóttir et al., 2013). The sites of the T-DNA cassette in both interesting mutants were also confirmed via PCR by amplifying fragments of T-DNA right and left border across right and left fungal flanking regions using two pairs of primers: g2420Compl-F2/HYGR and GFPF /Scf12-TDNA-R2 for strain PLM1983, and Scf61-TDNA-F1/HYGR and GFPF/Scf61-TDNA-R1 for strain PLM2739. Each PCR (25 µl) consisted of 5 µl 5xPCR buffer, 1 µl of 10mM dNTPs, 1 µl of 10mM each primer, 0.4 µl Taq, 1 µl DNA template, and 15.6 µl H_2O. Amplification conditions consisted an initiation step of 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 1 min and a final elongation step of 72 °C for 10 min. Furthermore, a part of hygromycin B was also amplified using pair primers HYGF/HYGR with the same condition above. Amplicons corresponding to the flanking regions of gDNA and T-DNA borders were sequenced via Sanger sequencing by Genewiz Inc. (South Plainfield, NJ, USA) to confirm the sites of T-DNA insertion in the mutants.

Updated prediction of gene g2420.t1 using RNA-seq

The gene g2420.t1 was queried with BLASTp v2.9.0 against the NCBI nr database (updated May 20, 2019), and the top five BLAST hits were mapped to the *D. longicolla* MSPL 10-6 genome (Li et al., 2015) with Exonerate v2.2.0 (Slater & Birney, 2005). *D. longicolla* RNA-seq reads (SRA accession SRX4349645) were mapped to the genome with GSNAP v2014-10-09 (Wu & Watanabe, 2005) and processed with SAMtools v1.7 (Li et al., 2009). The alignment of proteins and RNA-seq reads were visualized with IGV v2.4.16 (Robinson et al.,

2011). Furthermore, gene order and orientation of other predicted cytochrome P450s in scaffold 12 (Darwish et al., 2016b) were determined for insight about putative cytochrome P450 cluster genes related to the gene g2420.t1.

The tertiary structure of gene g2420.t1

After updating the prediction of the gene g2420.t1, tertiary structure of the protein of g2420.t1(cytochrome p450) in *D. longicolla* was predicted by the I-TASSER server (Yang & Zhang, 2015; Zhang et al., 2017).

Relative expression of gene g2420.t1

Quantitative reverse transcription PCR (qPCR) was utilized to measure the relative expression of g2420.t1 in strains PL2010 and PLM1983 at 0, 1, 3, and 5 days after inoculation. For each strain, four soybean stem segments (0.9 g) were inoculated by inserting 0.2 g of mycelia from strain PL2010 or PLM1983 into a vertically cut wound in stem segments using a sharp blade, followed by incubation in sterilized moisture conditions. Inoculated stem tissues were ground to a fine powder with liquid nitrogen. Total RNA was extracted with Ribozol according to the manufacturer's protocol (Amresco, Solon, OH, USA). For each sample, 2 µg total RNA was treated with DNase (Promega, Madison, WI, USA) for 30 min at 37 °C to remove genomic DNA. One µg of DNA-free RNA was utilized to synthesize cDNA with NxGen®M-MuLV reverse transcriptase following the manufacturer's protocol (Lucigen Corporation, Middleton, WI, USA). The qPCR reaction (10 µl) consisted of the following: 5 µl PerfeCTa™SYBR ® Green FastMix™ 2 × Master Mix (Quanta Biosciences, Inc., Gaithersburg, MD, USA), 250 nM of each forward and reverse primer, and 4 µl of cDNA template diluted 1:50 in nuclease-free water. qPCR was performed in the CFX Connect™ Real-Time PCR Detection

System (Bio-Rad Laboratories, Inc., Hercules, CA, USA). The primer set g2420qtrexon3-4F/g2420qtrexon3-4R for g2420.t1 and PlTubF1/PlTubR1 for beta-tubulin were designed with PrimerQuest (Integrated DNA Technologies). PCR conditions consisted of 1 cycle of 10 min at 95 °C, and 40 cycles of 15 s at 95 °C and 1 min at 58 °C. Expression data were collected with CFX Maestro™ Software (version1.1, Bio-Rad Laboratories, Inc., Hercules, CA, USA). Relative expression of g2420.t1 was normalized with beta-tubulin as the reference gene and calculated as fold changes (2^ΔΔCq) in expression relative to expression in the wild type as described by Taylor et al. (2019). Three technical replicates of qPCR were performed for each sample.

Complementation of *D. longicolla* mutants PLM1983 and PLM2739

Amplifying selected genes, Gibson assembly, and vector construction: To complement gene g2420.t1 in strain PLM1983 and gene g4126.t1 in strain PLM2739, genomic DNA of wild-type PL2010 *D. longicolla* was extracted with DNeasy Plant Mini Kit (Qiagen, MD, USA). A 6403 bp amplicon corresponding to gene g2420.t1 and a 2729 bp amplicon corresponding to gene g4126.t1 were amplified with g2420compl-F2/g2420compl-R2 for g2420.t1 and g4126compl-F3/g4126compl-R3 for g4126.t1. PCR reactions (50 µl) consisted of 5x LongAmp Taq Reaction Buffer (New England Biolabs) (10 µl), 10 mM dNTPs (1.5 µl), 10 mM of each forward or reverse primer (2 µl), LongAmp Taq DNA Polymerase (New England Biolabs) (1.5 µl), DNA template (2.5 µl) and sterile H_2O (30.5 µl). Amplification conditions included an initiation step (95 °C for 30 s), 35 amplification cycles (95 °C for 30 s, 54 °C for 30 s, and 65 °C for 210 s for g4126.t1 and 330 s for g2420.t1) and a final elongation step (65 °C for 10 min). PCR products were purified with Gene JET purification columns following the manufacturer's recommendations (Gel Extraction and DNA Cleanup Micro Kit; Thermo Fisher Scientific, Waltham, MA, USA). DNA was eluted in 35 µl H_2O and stored at 4 °C. The vector pBYR48 was

used as a backbone for complementation constructs (pBFA2420 and pBFA4126) after removing GFP via double digestion with BamHI and HindIII. Plasmid pBYR48 was originally created from pBYR14 by replacing hygromycin B_phosphotransferase (HYGR) with geneticin using SpeI/HindIII. To construct pBFA2420 and pBFA4126, a Gibson assembly reaction (homemade) was performed to fuse the pBYR48 backbone with purified PCR products of each gene (g2420.t1 and g4126.t1). Each reaction (12 µl) consisted of 1x Gibson master mix, 0.1 pmol of pBYR48 backbone (≈ 50 ng), and 0.2 pmol of each gene fragment was incubated in a thermocycler at 50°C for 60 min. Following incubation, samples were stored on ice for subsequent transformation. The complementation construct for g4126 was designated pBFA4126, and the complementation construct for g2420.t1 was designated pBFA2420 (Figure 1C, D).

***E. coli* and *A. tumefaciens* transformation**: To amplify complementation vectors assembled as described above, 50 µl of Mix and Go chemically competent *E. coli* cells (strain DH5α; prepared following manufacturer's instructions) (Zymo Research, Irvine, CA) were thawed on ice and mixed gently for 5 s using a micropipette with 5 µl of Gibson assembly product for each complementation construct. After transformation mixtures were incubated on ice for 5 min, 4 volumes of LB were added and incubated for 1 h at 37 °C with gentle shaking at 200 - 300 rpm. The mixtures were centrifuged, and 200 µl of supernatant was discarded. The resuspended cell mixture (10 µl) was spread onto pre-warmed LB plates containing kanamycin (50 µg/ml) and incubated overnight at 37 °C. The next day, 14 bacterial colonies were screened for complementation constructs via PCR amplification with primers Scf61-TDNA-F1/Scf61-TDNA-R1 for g4126 and g2420Compl-F2/Scf12-TDNA-R1. Each PCR reaction (25 µl) consisted of 5x PCR buffer (5 µl), 10 mM dNTPs (1 µl), 10 mM of each primer (1 µl), Taq polymerase (0.4 µl), 1 µl DNA template, and sterile H_2O (15.6 µl). Amplification conditions included an initiation

step (95 °C for 3 min), 35 amplification cycles (94 °C for 30 s, 56 °C for 30 s, and 72 °C for 1.5 min for g4126.t1 and 1 min for g2420.t1), and a final elongation step (72 °C for 10 min). Plasmid DNA was extracted from three *E. coli* colonies tested positive via PCR with a Zyppy™ Plasmid Miniprep Kit (Zymo Research, Irvine, CA). For confirmation via diagnostic enzymatic digestion (following the manufacturer's protocols), pBFA4126 was digested with EcoRI, HindIII, and PvuII; pBFA2420 was digested with BstEII, BamHI, and SphI (New England BioLabs, MA, USA). For the transformation of *A. tumefaciens*, electrocompetent *A. tumefaciens* cells were prepared following the slightly modified protocol of Wise et al. (2006). For transformation, plasmid DNA (~ 100 ng in 3 µl of water) was added to 50 µl of electrocompetent *A. tumefaciens* cells, mixed, and transferred to a chilled cuvette with a 0.2 cm gap. LB (1 ml) was added to the cuvette, mixed, and transferred to a 15 ml conical tube. The tube was incubated at 28 °C for 2 - 4h with shaking at 150 rpm. The mixtures were centrifuged and 890 µl of supernatant was discarded. The resuspended mixture (10 µl) was spread onto pre-warmed LB plates containing kanamycin (50 µg/ml) and carbenicillin (50 µg/ml) and incubated at 28 °C for 36 h. A total of 14 *A. tumefaciens* colonies were screened for the presence of complementation cassettes with primers Scf61-TDNA-F1/Scf61-TDNA-R1 for g4126.t1 and g2420Compl-F2/Scf12-TDNA-R1 for g2420.t1. PCR amplifications (25 µl) consisted of 5x PCR buffer (5 µl), 10 mM dNTPs (1 µl), 10 mM of each primer (1 µl), Taq polymerase (0.4 µl), 1 µl DNA template, and sterile H_2O (15.6 µl). Amplification conditions included an initiation step (95 °C for 3 min), 35 amplification cycles (94 °C for 30 s, 56 °C for 30 s, and 72 °C for 1:30 min for g4126.t1 or 1 min for g2420.t1), and a final elongation step (72 °C for 10 min). DNA from colonies that tested positive via PCR was extracted with a Zyppy™ Plasmid Miniprep Kit (Zymo Research, Irvine, CA) for confirmation via diagnostic enzymatic digestion. EcoRI, HindIII, and PvuII were used for

pBFA4126, and BstEII, BamHI, and SphI were used for pBFA2420 following the manufacturer's recommendations (New England Biolabs, MA, USA).

Genetic transformation was performed as described by Li et al. (2013) except geneticin (150 µg/ml) was used for selection. Transformants were screened for pycnidia production on oatmeal agar, and transformants with restored phenotype were evaluated for pathogenicity with the cut-stem assay. The complementation cassette was amplified by PCR with two primer sets, g2420-right-f/scf12-right-R and scf12-R/gen418-F for g2420.t1 and two primer sets g4126-right-f/g4126-right-R and gen418F/scf61R for g4126.t1. Amplification conditions included an initiation step (95 °C for 5 min), 35 amplification cycles (94 °C for 30 s, 55 °C for 50 s, and 72 °C for 1 min) and a final elongation step (72 °C for 10 min). PCR products were evaluated on 1% agarose gels and sequenced by Genewiz Inc. (South Plainfield, NJ, USA) to confirm the presence of reintroduced genes.

Statistical analysis

Data from the evaluation of growth and conidiation on different media, cut-stem pathogenicity assays, and gene expression were statistically compared by ANOVA using JMP pro14 software (SAS Institute Inc) complemented by Tukey's test. The means and standard error were calculated, and the different letters indicate statistical differences ($P \leq 0.05$).

Results

Fungal transformation and screening randomly tagged *D. longicolla* mutants

Agrobacterium tumefaciens mediated transformation generated 1,251 insertional mutant strains constitutively expressing GFP. These mutants were created by five separate transformations events. On average, each transformation produced 250 transformants. Two

rounds of screening to identify mutants with reduced seed colonization identified five mutants that were distinguished with a 0-7 scale according to Xue's scale (0-10) compared with wild-type. Zero represented seeds with no observed fungal growth and ten represented seeds completely covered by mycelium (Figures 2 and 3).

Evaluation of virulence in selected mutants

Impaired pathogenicity was confirmed for selected mutants with a cut seedling pathogenicity assay (Li et al., 2010). Statistically smaller necrotic lesions were induced by mutants on soybean stems compared to the wild type (Figures 4 and 5). Mutants PLM1868, PLM1983, and PLM2739 were highly impaired in the induction of necrosis on soybean stems, which indicated that their disrupted genes could be involved in colonization and/or pathogenesis in *D. longicolla*.

Evaluation of growth and sporulation in selected mutants

To evaluate growth and conidiation in selected mutants, strains were grown on four different media with wild-type strain PL2010 included as a positive control. Radial growth of PLM1983 was not reduced on any medium, and growth of PLM2080 was similar to the wild-type on three different media. However, growth of the other mutants was affected negatively on most media (Figure 6). Moreover, three of the five selected mutants were deficient in the formation of pycnidia and conidia, including PLM1868, PLM1983, and PL2739 (Figure 7).

Identification of genes disrupted in mutants PLM1983 & PLM2739

Target enrichment sequencing identified a single site of *Agrobacterium*-mediated T-DNA insertion in the mutants PLM1983 (Figures 8 and 9) and PLM2739 (Figures 11 and 12). T-DNA insertion of the mutant PLM1983 was identified in scaffold 12 of the genome, approximately

1564 bp from the predicted start codon of g2420.t1 (predicted to encode a cytochrome P450) (Figures 8, 9, and 10). Read coverage was 4003x for the insertion site closest to the coding region of the gene (Figure 9). On the other hand, the T-DNA insertion in mutant PLM2739 disrupted the coding region of an unknown gene, g4126.t1, in scaffold 61 that shared low levels of similarity with a putative serine threonine protein (Figures 11, 12, and 13). The maximum coverage was 18,295 reads in one of the flanking regions of insertion T-DNA (Figure 13).

According to gene prediction for the whole genome of *D. longicolla* by Darwish et al. (2016b), 24 predicted cytochrome P450 genes including g2420.t1 in scaffold 12 were determined with their order and orientation. Some of them could be involved in pathogenicity and asexual reproduction as putative cytochrome P450 cluster genes related to the gene g2420.t1 (Figure 14).

Updated annotation of g2420.t1 via RNA-seq

Updated prediction of g2420.t1 based on RNA-seq coverage indicates the gene is shorter (1,885 bp) than predicted previously (3,568 bp) (Darwish et al., 2016a) indicating a possible error in the previous annotation (Figure 15). Additionally, homologous cytochrome p450 proteins (GenBank accessions POS69387.1; ROW02277.1; ROW09819.1; ROW14554.1; KUI68930.1; P54781) generally have only CYPX domain. However, the previous annotation of g2420.t1 predicted two domains, which are CYPX and Rhodanese-like domain. Therefore, the updated prediction of this gene is likely to be more accurate than the initial prediction.

Predicted tertiary structure of g2420.t1 (cytochrome P450)

The tertiary structure of the protein encoded by g2420.t1 in *D. longicolla* was predicted using the I-TASSER server (Yang & Zhang, 2015; Zhang et al., 2017). The prediction identified secondary structure elements of the tertiary structure of g2420.t1 protein, represented as a-helices

in red, b-sheets in yellow, and loops in blue. The heme is represented in ball and sticks (Figure 16).

Relative expression of g2420.t1

The relative expression of g2420.t1 in the wild-type strain PL2010 of *D. longicolla* and the mutant PLM1983 via RT-qPCR after 0, 1, 3, and 5 days post-inoculation showed reduced expression in strain PLM1983 (Figure 17). However, no significant difference was observed regarding time scales, which suggests that the gene may not be induced during early stages of pathogenesis.

Complementation of *D. longicolla* mutants PLM1983 and PLM2739

To confirm the potential role of genes identified by forward genetic screening in virulence and conidiation, mutants were genetically complemented. Twenty-six complemented transformants of mutant PLM1983 were obtained from three transformation events. Two transformants partially restored the wild-type phenotype by producing pycnidia (Figure 18). Like the wild-type, one of these transformants also induced necrotic lesions on soybean stems, albeit not fully at the wild-type level (Figure 19). To complement the mutant PLM2739, 200 transformants were created. However, none of these transformants restored the wild-type phenotype (Figure 20). This lack of complementation suggests there is no link between the gene and the phenotype, or possibly a defect within the complementation construct. The presence of complementation cassettes in the mutants was confirmed by PCR with two primer sets, g2420-right-f/scf12-right-R and scf12-R/gen418-F for g2420.t1 and two primer sets g4126-right-f/g4126-right-R and gen418F/scf61R for g4126.t1, indicating that the genes were successfully reintroduced into the mutants (Figure 21).

Discussion

Agrobacterium tumefaciens-mediated transformation (ATMT) was successfully used in a previous study to perform mutagenesis in *D. longicolla* with an average of 150 - 250 transformants per 1 transformation (Li et al., 2013). Similarly, transformations in this study yielded an average of 250 transformants per transformation. After screening 1,251 transformants, five mutants were visually defective in colonizing soybean seeds. This reduction in external colonization by mutants of *D. longicolla* could correlate positively with the internal colonization of seed by these mutants. Xue et al. (2006) observed a linear relationship between a visual assessment of soybean seed colonization by *D. phaseolorum* and *C. kikuchii* and ergosterol content, which is a fungal-specific membrane sterol. Furthermore, mycelium dry masses of these fungi in seeds have a strong linear relationship with ergosterol content. Results of cut-stem pathogenicity assays also indicated that mutants were impaired in inducing necrotic lesions compared to the wild type. Despite the fact that soybean stems were cut and physical barriers for infection were removed, these mutants could not induce extensive necrosis on stems.

The mutant PLM1983 was an interesting strain due to its impairment in seed colonization and stem necrosis. Furthermore, the disruption did not reduce radial growth on culture media although it affected mycelial density. Target enrichment sequencing successfully identified one insertion site of mutant PLM1983 in the upstream region of g2420.t1, which was predicted to encode a cytochrome P450 ortholog. This disruption within the putative promoter region of g2420.t1 could impair transcription of the gene and consequently gene expression (Figure 17). Comparative analyses of g2420.t1 orthologs in other fungi, g2420.t1 belongs to the CYP61 clan. The CYP61 family has a unique motif (ASQDAS/T) that distinguishes it from the other 15 cytochrome P450 clades (Chen et al., 2014).

The function of gene g2420.t1 could be linked to fundamental components of seed colonization, inducing the induction of necrotic lesions, as well as the production of conidia. Compared to plants and animals, relatively few members of the cytochrome P450 superfamily have been functionally characterized in fungi (Shin et al., 2018). Functionally characterized examples have roles in pathogenicity (Wang et al., 2019), detoxification and production of mycotoxins (Crešnar & Petrič, 2011), and ergosterol biosynthesis (Skaggs et al., 1996). The link between reduction of necrosis and colonization, and disruption of a cytochrome p450, could implicate g2420.t1 in toxin biosynthesis. Several fungal cytochrome p450s have been implicated in toxin biosynthesis, such as aflatoxin (Yu et al., 1997) and trichothecenes (Cardoza et al., 2011). Furthermore, treatments of soybean seedlings with culture filtrates of *D. longicolla* and *D. sojae* caused wilting and necrosis of excised soybean seedlings. Also, culture filtrates of *D. longicolla* significantly inhibit soybean seedling radicle growth (Ivanovic & Sinclair, 1989). *D. longicolla* can behave as a necrotrophic pathogen and kill the plant tissues before colonizing them (Ivanovic & Sinclair, 1989). On the other hand, Kung et al., (1976) also indicated that feed amended with *Phomopsis* sp. caused hepatic necrosis with high mortality in chicks because this fungus could produce mycotoxins that increase the level of liver glucose-6-P dehydrogenase.

Gene g2420.t1 may be a member of an uncharacterized cytochrome p450 cluster in *D. longicolla*. The genome annotation of *D. longicolla* predicted 343 members of the cytochrome p450 superfamily (Darwish et al., 2016b). Furthermore, 24 CYPs were predicted in scaffold 12, including g2420.t1. These genes could potentially behave as a cluster and be regulated by transcriptional factors in the same regulatory pathway. Additionally, the relatively large number of CYPs predicted in the *D. longicolla* genome could indicate a widespread role of these genes in pathogenesis. In other fungi, the number of CYPs positively correlated with pathogenicity,

survival, and niche adaptation. Filamentous fungi generally have more CYPs compared to relatively few CYPs of yeast-like fungi. However, plant pathogenic fungi tend to have large numbers of CYP genes (Shin et al., 2018).

Further investigation will be required to identify the exact pathway and potential gene cluster associated with g2420.t1. However, these results of this study suggest a critical role for g2420.t1 in pathogenicity and asexual reproduction. Targeting this gene and its cluster may provide a novel strategy for controlling *Diaporthe* diseases of soybean.

Literature Cited

Agrios, G. N. (2005). *Plant pathology*: Edition 5th Academic Press, New York, United States of America. 922 p.

Allen, T. W., Bissonnette, K., Bradley, C. A., Damicone, J. P., Dufault, N. S., Faske, T. R., . . . Young, H. (2018). *Southern United States soybean disease loss estimates for 2017*. Paper presented at the The 45th Annual Meeting of the Southern Soybean Disease Workers, March 7–8, Pensacola Beach, Florida.

Allen, T. W., Bissonnette, K., Bradley, C. A., Damicone, J. P., Dufault, N. S., Faske, T. R., . . . Young, H. (2019). *Southern United States soybean disease loss estimates for 2018.* Paper presented at the The 46 th Annual Meeting of the Southern Soybean Disease Workers, March 6–7, Pensacola Beach, Florida.

Baker, D. M., Minor, H. C., Brown, M. F., & Brown, E. A. (1987). Infection of immature soybean pods and seeds by *Phomopsis longicolla*. *Canadian Journal of Microbiology,* 33(9), 797-801. doi:10.1139/m87-135

Cardoza, R. E., Malmierca, M. G., Hermosa, M. R., Alexander, N. J., McCormick, S. P., Proctor, R. H., . . . Gutiérrez, S. (2011). Identification of loci and functional characterization of trichothecene biosynthesis genes in filamentous fungi of the genus Trichoderma. *Appl. Environ. Microbiol*., 77(14), 4867-4877.

Chen, W., Lee, M.-K., Jefcoate, C., Kim, S.-C., Chen, F., & Yu, J.-H. (2014). Fungal cytochrome p450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin. *Genome Biology and Evolution,* 6(7), 1620-1634.

Crešnar, B., & Petrič, Š. (2011). Cytochrome P450 enzymes in the fungal kingdom. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics,* 1814(1), 29-35.

Cui, Y. L., Duan, C. X., Wang, X. M., Li, H. J., & Zhu, Z. D. (2009). First report of *Phomopsis longicolla* causing soybean stem blight in China. *Plant Pathology,* 58(4), 799-799. doi:10.1111/j.1365-3059.2009.02057.x

Córdova, P., Gonzalez, A. M., Nelson, D. R., Gutiérrez, M. S., Baeza, M., Cifuentes, V., & Alcaíno, J. (2017). Characterization of the cytochrome P450 monooxygenase genes (P450ome) from the carotenogenic yeast *Xanthophyllomyces dendrorhous*. *BMC Genomics,* 18(1), 540. doi:10.1186/s12864-017-3942-9

Darwish, O., Li, S. X., Matthews, B., & Alkharouf, N. (2016a). Genome-wide functional annotation of *Phomopsis longicolla* isolate MSPL 10-6. *Genomics Data*, 8, 67-69. doi:10.1016/j.gdata.2016.03.006

Darwish, O., Li, S. X., May, Z., Matthews, B., & Alkharouf, N. W. (2016b). A searchable database for the genome of *Phomopsis longicolla* (isolate MSPL 10-6). *Bioinformation,* 12(4), 233-236. doi:10.6026/97320630012233

Gillen, A. M., Smith, J. R., Mengistu, A., & Bellaloui, N. (2012). Effects of maturity and *Phomopsis longicolla* on germination and vigor of soybean seed of near-isogenic lines. *Crop Science,* 52(6), 2757-2766. doi:10.2135/cropsci2011.10.0566

Hepperly, P. R., & Sinclair, J. B. (1980). Detached pods for studies of *Phomopsis sojae* pods and seed colonization. *The Journal of Agriculture of the University of Puerto Rico,* 64(3), 330-337.

Isaka, M., Jaturapat, A., Rukseree, K., Danwisetkanjana, K., Tanticharoen, M., & Thebtaranonth, Y. (2001). Phomoxanthones A and B, novel xanthone dimers from the endophytic fungus *Phomopsis* species. *Journal of Natural Products,* 64(8), 1015-1018.

Ivanovic, M., & Sinclair, J. B. (1989). Comparison of possible phytotoxic metabolites in culture filtrates of the *Diaporthe Phomopsis* complex of soybeans. *Mycopathologia,* 108(1), 59-63. doi:10.1007/bf00436785

Jackson, E. W., Feng, C. D., Fenn, P., & Chen, P. Y. (2009). Genetic mapping of resistance to Phomopsis Seed Decay in the soybean breeding line MO/PSD-0259 (PI562694) and plant introduction 80837. *Journal of Heredity,* 100(6), 777-783. doi:10.1093/jhered/esp042

Kmetz, K. T., Schmitthenner, A. F., & Ellett, C. W. (1978). Soybean seed decay: Prevalence of infection and symptom expression caused by *Phomopsis* sp., *Diaporthe phaseolorum* var. *sojae*, and *D. phaseolorum* var. *caulivora*. *Phytopathology,* 68(6), 836-841.

Kulik, M. M., & Sinclair, J. B. (1999). Phomopsis seed decay. In G. L. Hartman, J. B. Sinclair, & J. C. Rupe (Eds.), *Compendium of soybean diseases,* (4 ed., pp.31-32). St. Paul, MN: American Phytopathological Society.

Kung, H. C., Chipley, J. R., Latshaw, J. D., & Kerr, K. M. (1976). Mycotoxicosis in chicks produced by toxins from *Phomopsis* sp or *Diaporthe phaseolorum* var. *solae*. *Avian Diseases,* 20(3), 504-518. doi:10.2307/1589383

Lazo, G. R., Stein, P. A., & Ludwig, R. A. (1991). A DNA transformation–competent Arabidopsis genomic library in *Agrobacterium*. *Biotechnology*, 9(10), 963.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:*1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Subgroup, G. P. D. P. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* 25(16), 2078-2079. doi:10.1093/bioinformatics/btp352

Li, S. (2011). Phomopsis seed decay of soybean. In *Soybean-Molecular Aspects of Breeding* (Sudaric A, ed.). IntechOpen 277–92. Rijeka, Croatia

Li, S., Ridenour, J. B., Kim, H., Hirsch, R. L., Rupe, J. C., & Bluhm, B. H. (2013). *Agrobacterium tumefaciens*-mediated transformation of the soybean pathogen *Phomopsis longicolla*. *Journal of Microbiological Methods,* 92(3), 244-245.

Li, S., Rupe, J., Chen, P., Shannon, G., Wrather, A., & Boykin, D. (2015). Evaluation of diverse soybean germplasm for resistance to Phomopsis seed decay. *Plant Disease,* 99(11), 1517-1525.

Li, S., Sciumbato, G., Rupe, J., Shannon, G., Chen, P., & Boykin, D. (2017). Evaluation of commercial soybean cultivars for reaction to Phomopsis seed decay. *Plant Disease,* 101(12), 1990-1997.

Li, S. X., Darwish, O., Alkharouf, N., Matthews, B., Ji, P. S., Domier, L. L., . . . Bluhm, B. H. (2015). Draft genome sequence of *Phomopsis longicolla* isolate MSPL 10-6. *Genomics Data*, 3, 55-56. doi:10.1016/j.gdata.2014.11.007

Li, S. X., Hartman, G. L., & Boykin, D. L. (2010). Aggressiveness of *Phomopsis longicolla* and Other *Phomopsis* spp. on Soybean. *Plant Disease,* 94(8), 1035-1040. doi:10.1094/pdis-94-8-1035

Li, S. X., Musungu, B., Lightfoot, D., & Ji, P. S. (2018). The interactomic analysis reveals pathogenic protein networks in *Phomopsis longicolla* underlying seed decay of soybean. *Frontiers in Genetics,* 9, 104. doi:10.3389/fgene.2018.00104

Lin, C.-H., Yang, S. L., Wang, N.-Y., & Chung, K.-R. (2010). The FUS3 MAPK signaling pathway of the citrus pathogen *Alternaria alternata* functions independently or cooperatively with the fungal redox-responsive AP1 regulator for diverse developmental, physiological and pathogenic processes. *Fungal Genetics and Biology,* 47(4), 381-391.

Lorang, J. M., Tuori, R. P., Martinez, J. P., Sawyer, T. L., Redman, R. S., Rollins, J. A., . . . Dickman, M. B. (2001). Green fluorescent protein is lighting up fungal biology. *Appl. Environ. Microbiol.,* 67(5), 1987-1994.

McGee, D. C. (1984). Epidemiology of soybean seed decay by *Phomopsis* and *Diaporthe* spp. *Seed Science and Technology,* 11, 1-11.

Mengistu, A., Castlebury, L. A., Smith, J. R., Rossman, A. Y., & Reddy, K. N. (2007). Isolates of *Diaporthe-Phomopsis* from weeds and their effect on soybean. *Canadian Journal of Plant Pathology-Revue Canadienne De Phytopathologie,* 29(3), 283-289.

Mullins, E. D., Chen, X., Romaine, P., Raina, R., Geiser, D. M., & Kang, S. (2001). Agrobacterium-mediated transformation of *Fusarium oxysporum*: an efficient tool for insertional mutagenesis and gene transfer. *Phytopathology,* 91(2), 173-180.

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* 26(6), 841-842. doi:10.1093/bioinformatics/btq033

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, 29(1), 24.

Roy, K. W., & Abney, T. S. (1988). Colonization of pods and infection of seeds by *Phomopsis longicolla* in susceptible and resistant soybean lines inoculated in the greenhouse. *Canadian Journal of Plant Pathology,* 10(4), 317-320.

Schmitt, M. W., Fox, E. J., Prindle, M. J., Reid-Bayliss, K. S., True, L. D., Radich, J. P., & Loeb, L. A. (2015). Sequencing small genomic targets with high efficiency and extreme accuracy. *Nature Methods,* 12(5), 423-425. doi:10.1038/nmeth.3351

Schneeberger, K. (2014). Using next-generation sequencing to isolate mutant genes from forward genetic screens. *Nature Reviews Genetics,* 15(10), 662.

Shin, J., Kim, J. E., Lee, Y. W., & Son, H. (2018). Fungal Cytochrome P450s and the P450 Complement (CYPome) of *Fusarium graminearum*. *Toxins,* 10(3), 112. doi:10.3390/toxins10030112

Sinclair, J. B. (1993). Phomopsis seed decay of soybeans: A prototype for studying seed disease. *Plant Disease*, 77(4), pp.329-334.

Skaggs, B. A., Alexander, J. F., Pierson, C. A., Schweitzer, K. S., Chun, K. T., Koegel, C., . . . Bard, M. (1996). Cloning and characterization of the *Saccharomyces cerevisiae* C-22 sterol desaturase gene, encoding a second cytochrome P-450 involved in ergosterol biosynthesis. *Gene,* 169(1), 105-109. doi:10.1016/0378-1119(95)00770-9

Slater, G. S., & Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics,* 6, 31. doi:10.1186/1471-2105-6-31

Smith, S., Fenn, P., Chen, P. Y., & Jackson, E. (2008). Inheritance of resistance to Phomopsis Seed Decay in PI 360841 soybean. *Journal of Heredity,* 99(6), 588-592. doi:10.1093/jhered/esn037

Staben, C., Jensen, B., Singer, M., Pollock, J., Schechtman, M., Kinsey, J., & Selker, E. (1989). Use of a bacterial hygromycin B resistance gene as a dominant selectable marker in *Neurospora crassa* transformation. *Fungal Genetics Reports,* 36(1), 79.

Taylor, S. C., Nadeau, K., Abbasi, M., Lachance, C., Nguyen, M., & Fenrich, J. (2019). The Ultimate qPCR experiment: Producing publication quality, reproducible data the first time. *Trends Biotechnol,* 37(7), 761-774. doi:10.1016/j.tibtech.2018.12.002

TeKrony, D. M., Egli, D. B., Stuckey, R. E., & Loeffler, T. M. (1985). Effect of benomyl applications on soybean seedborne fungi, seed germination, and yield. *Plant Disease,* 69(9), 763-765.

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform,* 14(2), 178-192. doi:10.1093/bib/bbs017

Tolbert, A. C., & Spurlock, T. N. (2017). First report of *Phomopsis longicolla* causing stem necrosis on soybean in Arkansas. *Plant Disease,* 101(12), 2147-2147. doi:10.1094/pdis-02-17-0223-pdn

Wang, Y., Wu, Q., Liu, L., Li, X., Lin, A., & Li, C. (2019). MoMCP1, a Cytochrome P450 gene, is required for alleviating manganese toxin revealed by transcriptomics analysis in *Magnaporthe oryzae*. *International Journal of Molecular Sciences,* 20(7), 1590.

Wise, A. A., Liu, Z., & Binns, A. N. (2006). Three methods for the introduction of foreign DNA into *Agrobacterium*. *Methods Mol. Biol*. 343: 43–53

Wrather, J. A., Shannon, J. G., Stevens, W. E., Sleper, D. A., & Arelli, A. P. (2004). Soybean cultivar and foliar fungicide effects on *Phomopsis* sp. seed infection. *Plant Disease,* 88(7), 721-723.

Wu, T. D., & Watanabe, C. K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics,* 21(9), 1859-1875. doi:10.1093/bioinformatics/bti310

Xue, C. S., Lu, Y. Y., Xiao, S. Q., & Duan, Y. X. (2015). First report of *Phomopsis longicolla* causing leaf spot on soybean in China. *Plant Disease*, 99(2), 290-290.

Xue, H. Q., Upchurch, R. G., & Kwanyuen, P. (2006). Ergosterol as a quantifiable biomass marker for *Diaporthe haseolorum* and *Cercospora kikuchi*. *Plant Disease,* 90(11), 1395-1398.

Yang, J. Y., & Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research,* 43(W1), W174-W181. doi:10.1093/nar/gkv342

Yu, J., Chang, P.-K., Cary, J. W., Bhatnagar, D., & Cleveland, T. E. (1997). AvnA, a gene encoding a cytochrome P-450 monooxygenase, is involved in the conversion of averantin to averufin in aflatoxin biosynthesis in *Aspergillus parasiticus*. *Appl. Environ. Microbiol*., 63(4), 1349-1356.

Zaccaron, M. L. (2019). Studies on pathogenesis of the diseases caused by *Macrophomina phaseolina* and *Phomopsis longicolla* on Soybean. Ph.D. Dissertation, University of Arkansas.

Zhang, C. X., Freddolino, P. L., & Zhang, Y. (2017). COFACTOR: improved protein function prediction by combining structure, sequence and protein-protein interaction information. *Nucleic Acids Research,* 45(W1), W291-W299. doi:10.1093/nar/gkx366

Figures and Tables

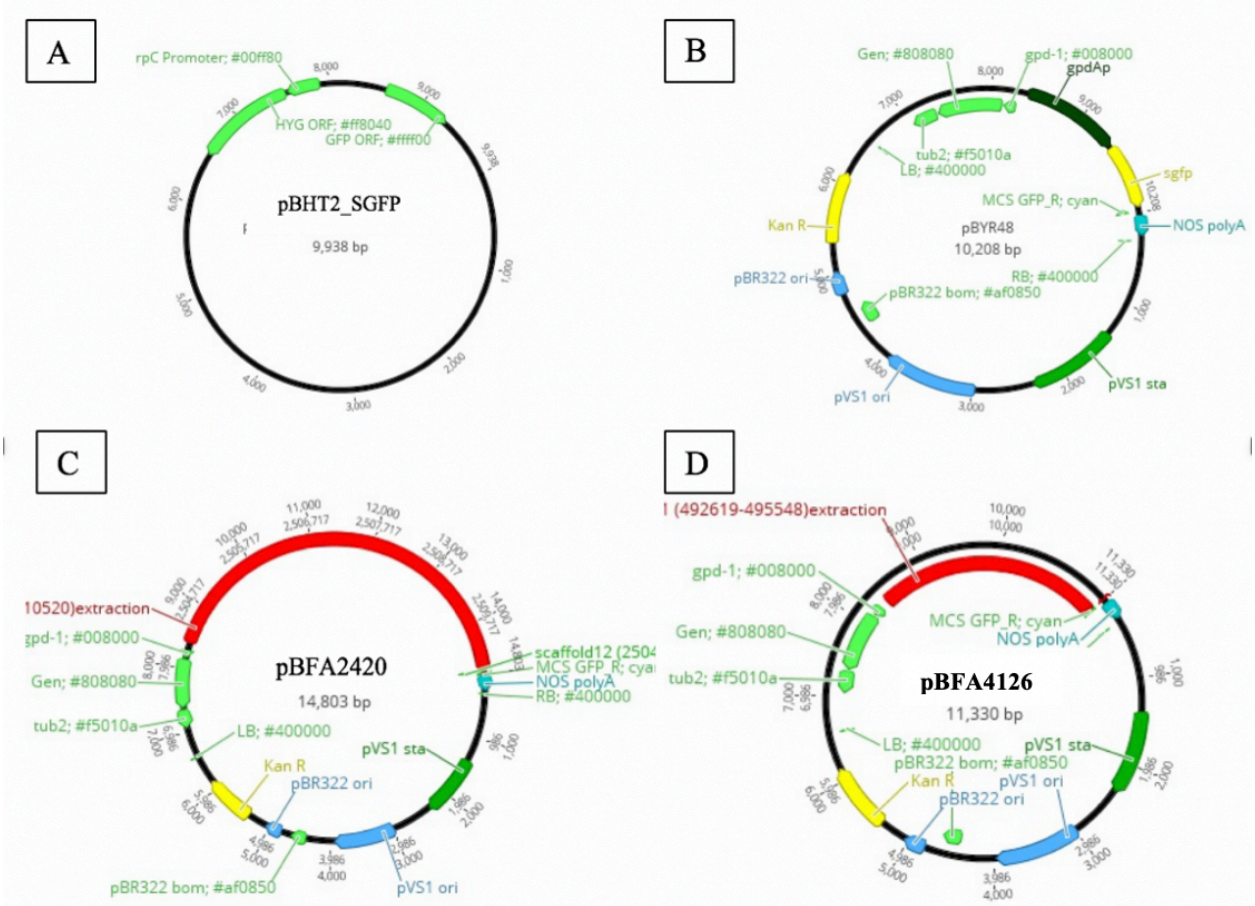


Figure 1. Plasmid maps of **(A)** pBH2_SGFP. **(B)** pBYR48. **(C)** pBFA2420. **(D)** pBFA4126.

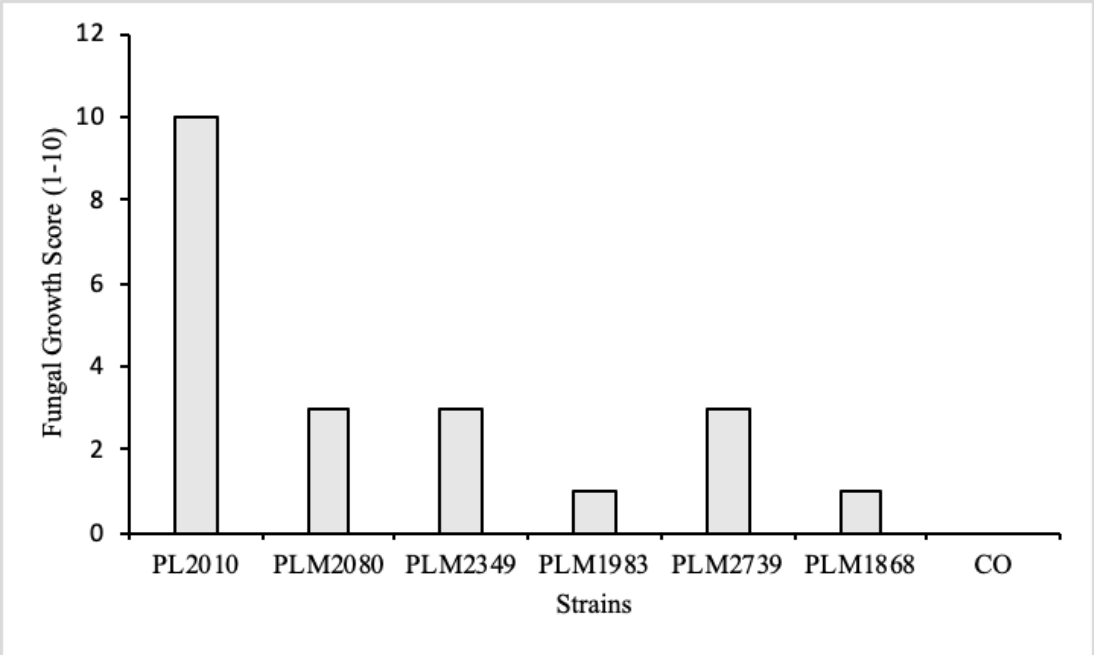


Figure 2. Fungal colonization on soybean seeds by five defected mutants and wild-type of *D. longicolla,* strain PL2010. Visual assessments of seed colonization were rated with a scale (0-10). Zero was represented seeds with no observed fungal growth and seeds completely covered by mycelium were scored ten (Xue et al., 2006).
Co = Control treatment without fungi

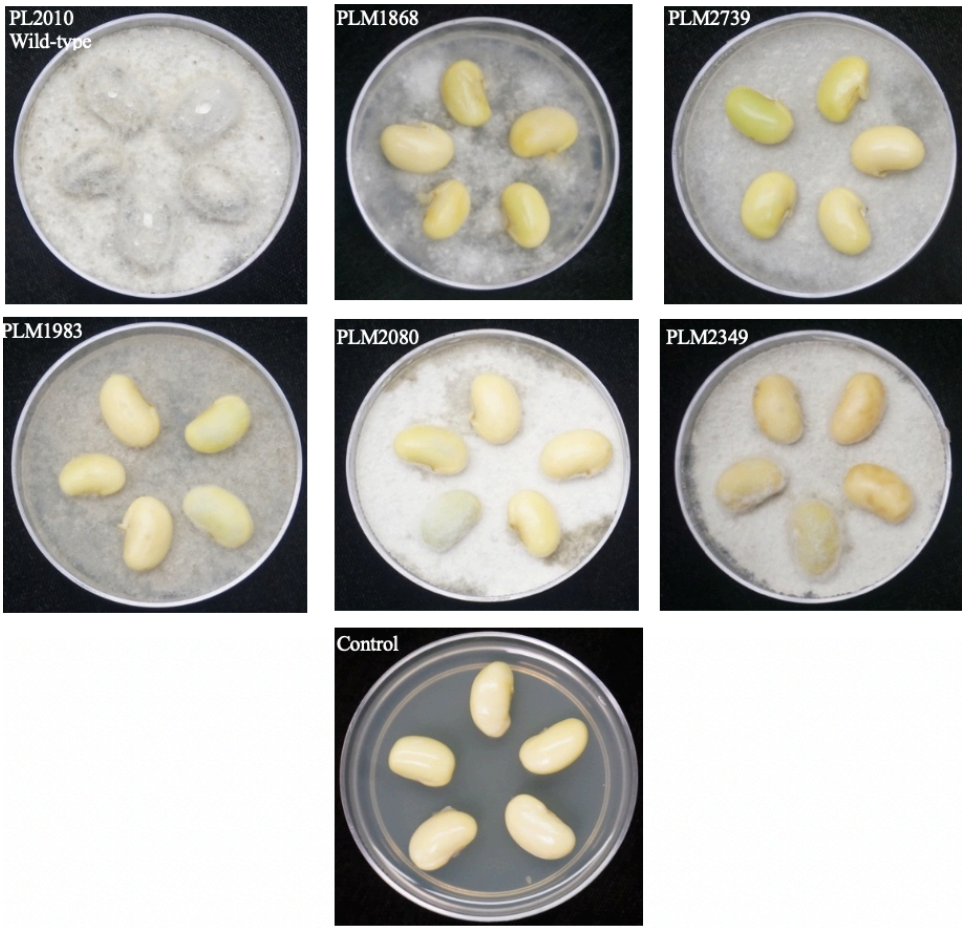


Figure 3. Colonization of soybean seeds by mutants and wild-type strain PL2010 of *D. longicolla*. Five seeds on the surface of mycelia growing in 6-cm petri dishes containing PDA. Cultures were incubated in continuous lighting at 25 °C for 5 days.

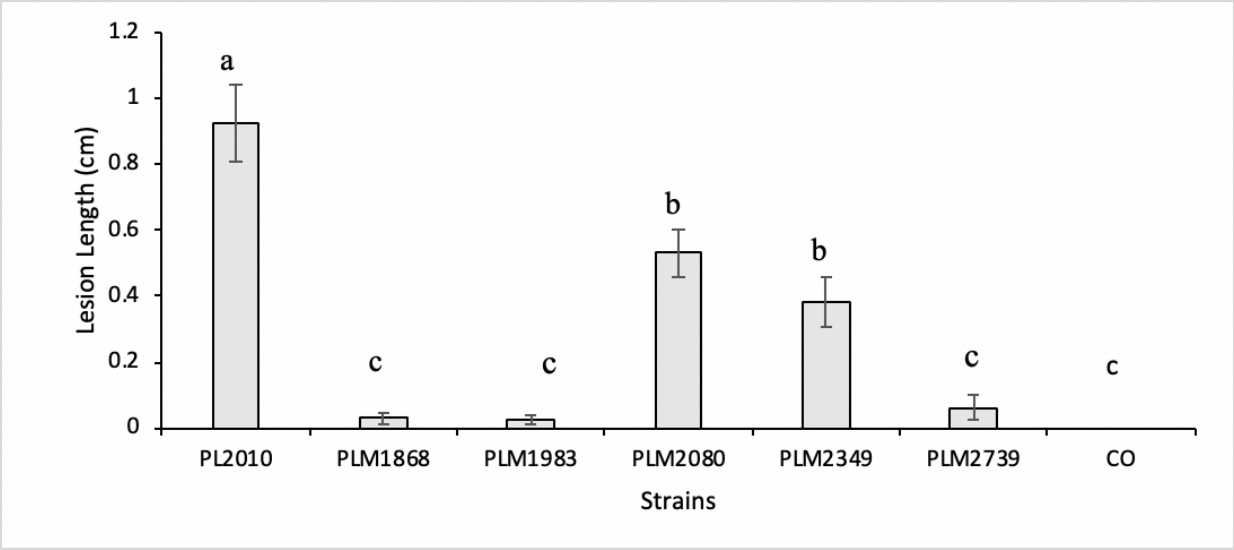


Figure 4. Necrotic lesions on soybean stems caused by mutants and wild-type strain PL2010 of *D. longicolla* with the cut-stem inoculation assay. Bars with different letters are significantly different according to the Tukey test ($p < 0.05$).

Figure 5. Necrotic lesions on soybean stems caused by mutants and wild-type strain PL2010 of *D. longicolla* with the cut-stem inoculation assay.

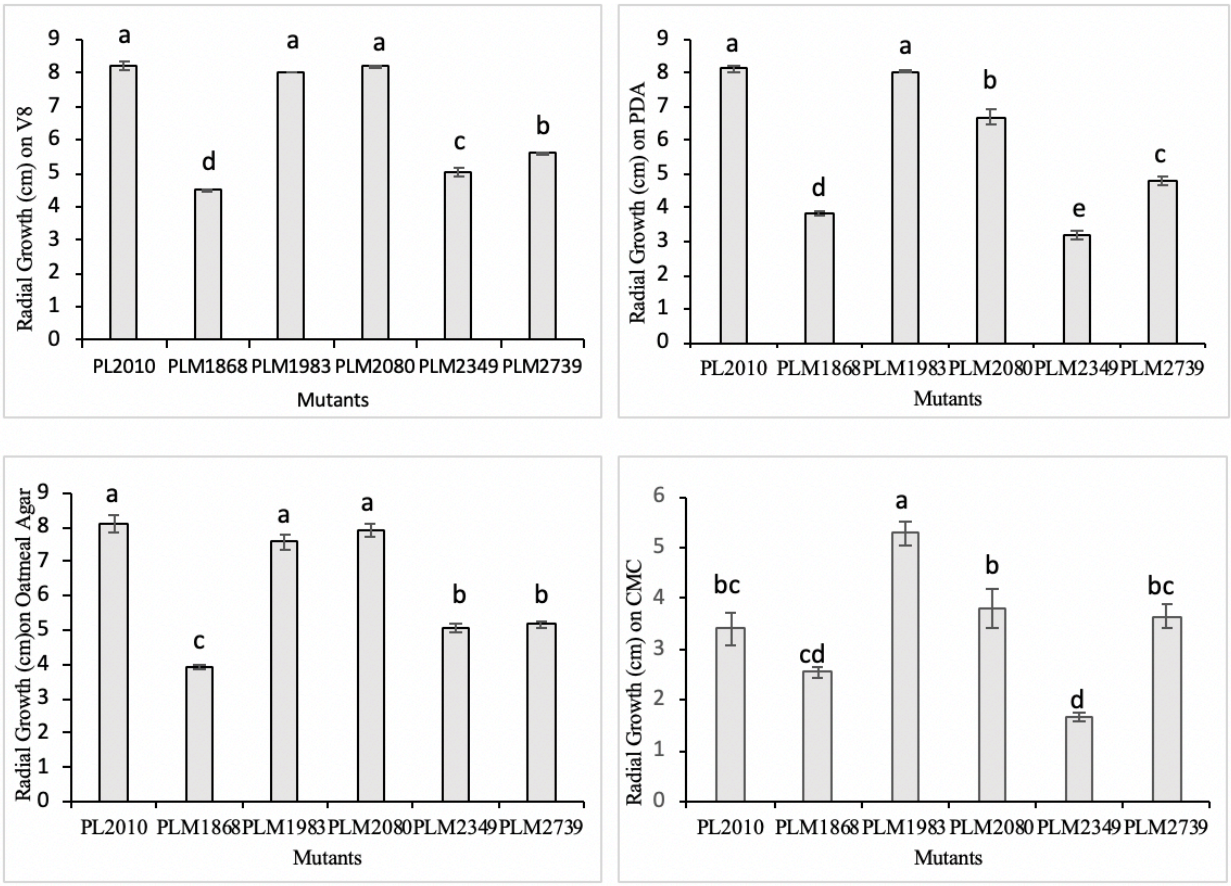


Figure 6. Radial growth of mutants and wild-type strain PL2010 of *D. longicolla* on four different media. Each value represents the average of four cultures for each strain. Bars with different letters are significantly different according to the Tukey test ($p \leq 0.05$).

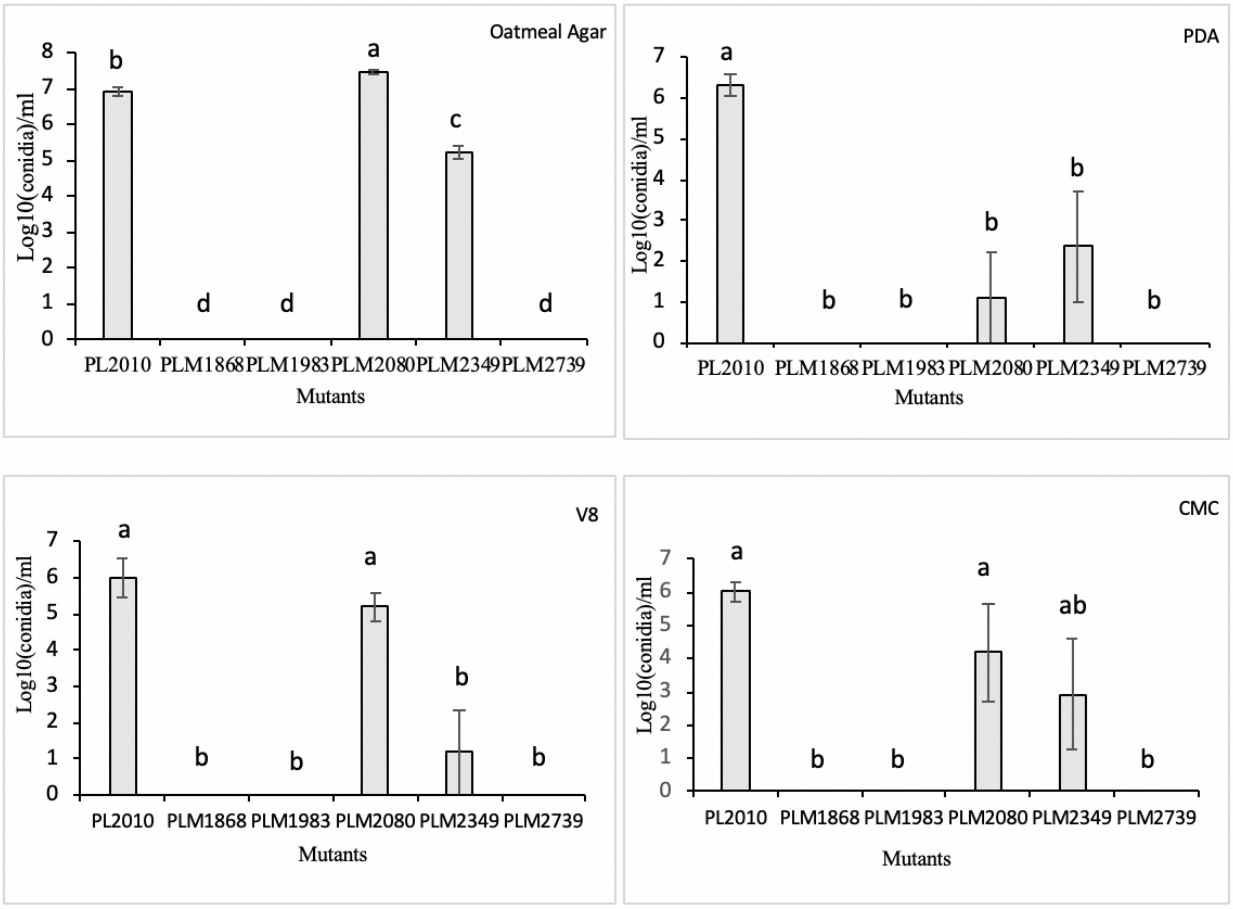


Figure 7. Sporulation of mutants and wild-type strain PL2010 of *D. longicolla* on different media. **(A)** on Oatmeal Agar. **(B)** on PDA. **(C)** on V8. **(D)** on CMC. Bars represent Log10 of conidia per ml. Each value represents the average of four replicates for each strain. Bars with different letters are significantly different according to the Tukey test ($p \leq 0.05$).

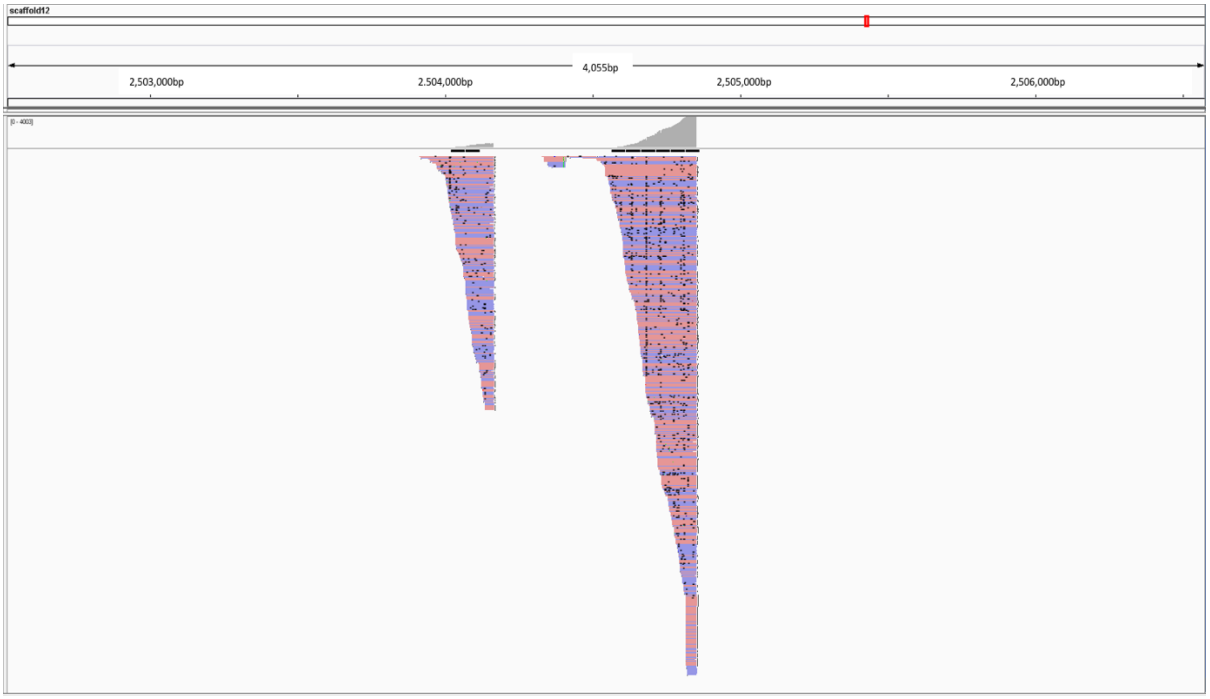


Figure 8. The location of inserted T-DNA in scaffold 12 in the genome of mutant PLM1983. Reads were mapped to the reference genome sequence of *P. longicolla* MSPL 10-6 (Li et al. 2015b).

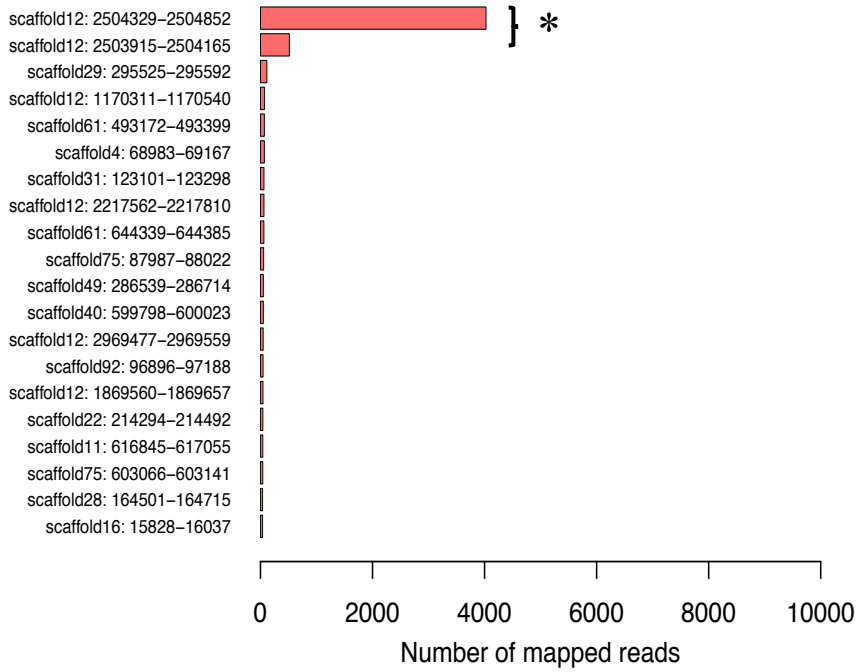


Figure 9. T-DNA read mapping of mutant PLM1983.

* Possible single insertion site.

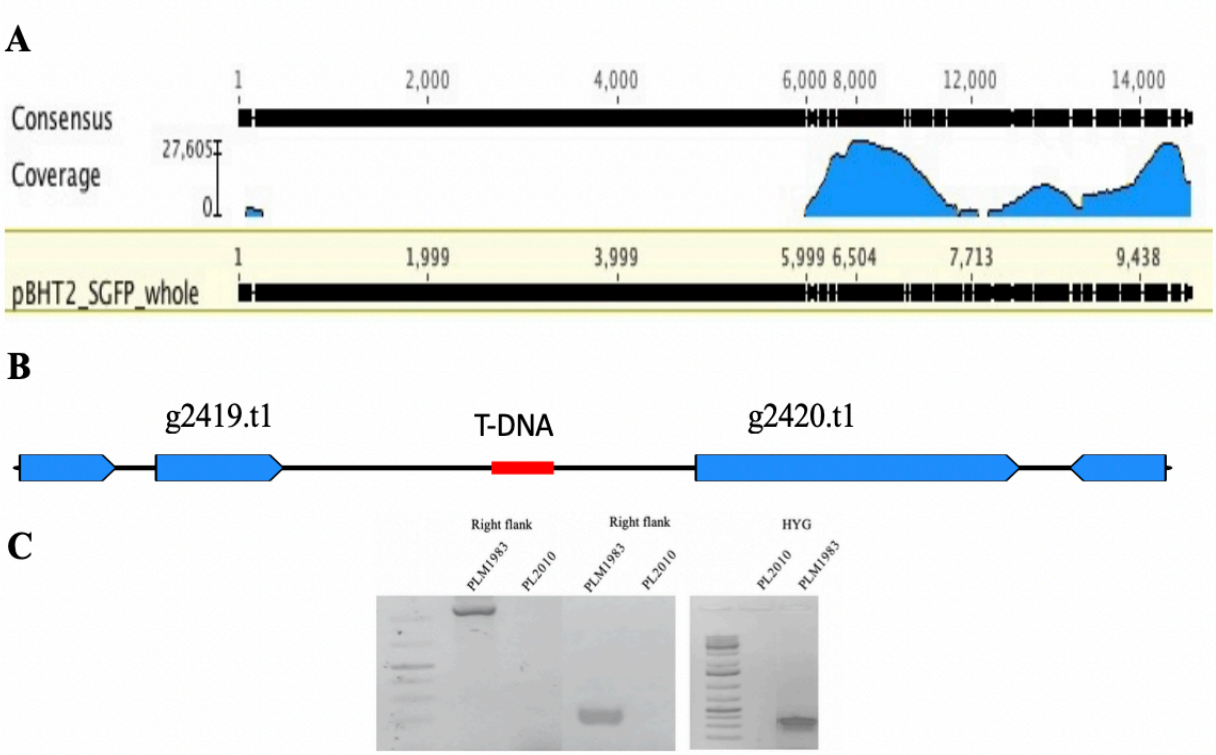


Figure 10. **(A)** Coverage and presence of reads of T-DNA in the mutant PLM1983 mapped to the plasmid, pBHt2_sGFP. **(B)** The site of T-DNA in the upstream region of the gene g2420.t1. **(C)** Bands of overlapping flanking regions of g2420.t1 with the inserted cassette (T-DNA) and a fragment of HYG compared with the wild-type strain, PL2010.

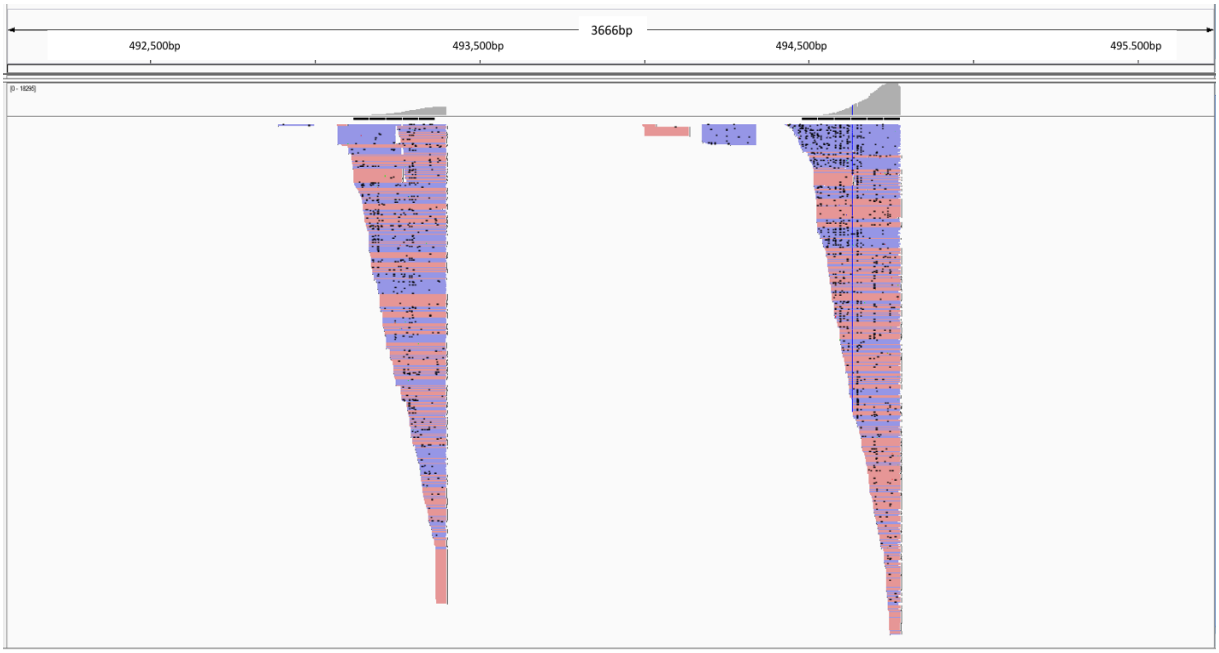


Figure 11. The location of inserted T-DNA in scaffold 61 in the genome of mutant PLM2739.

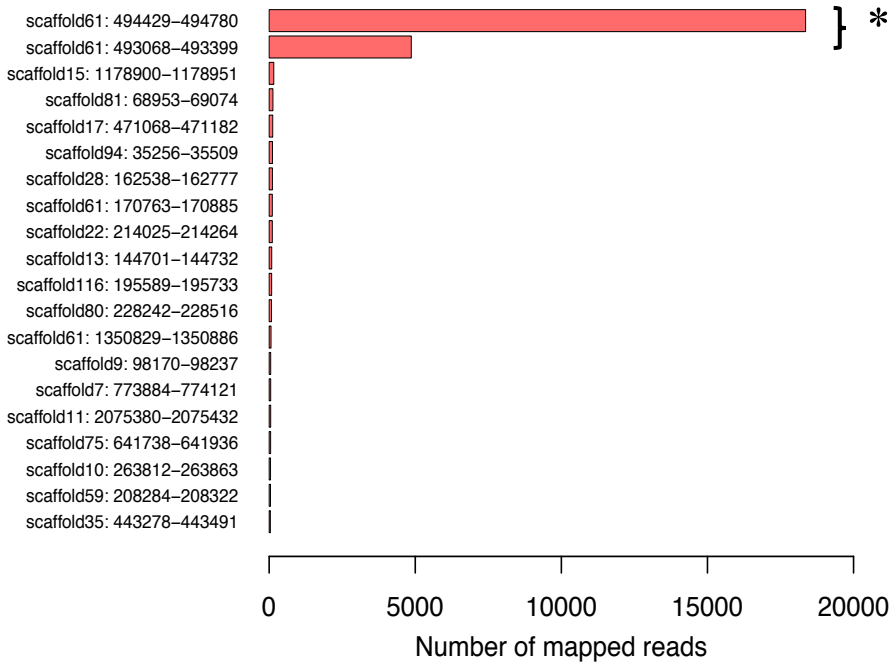


Figure 12. T-DNA read mapping of mutant PLM2739.

* Possible single insertion site.

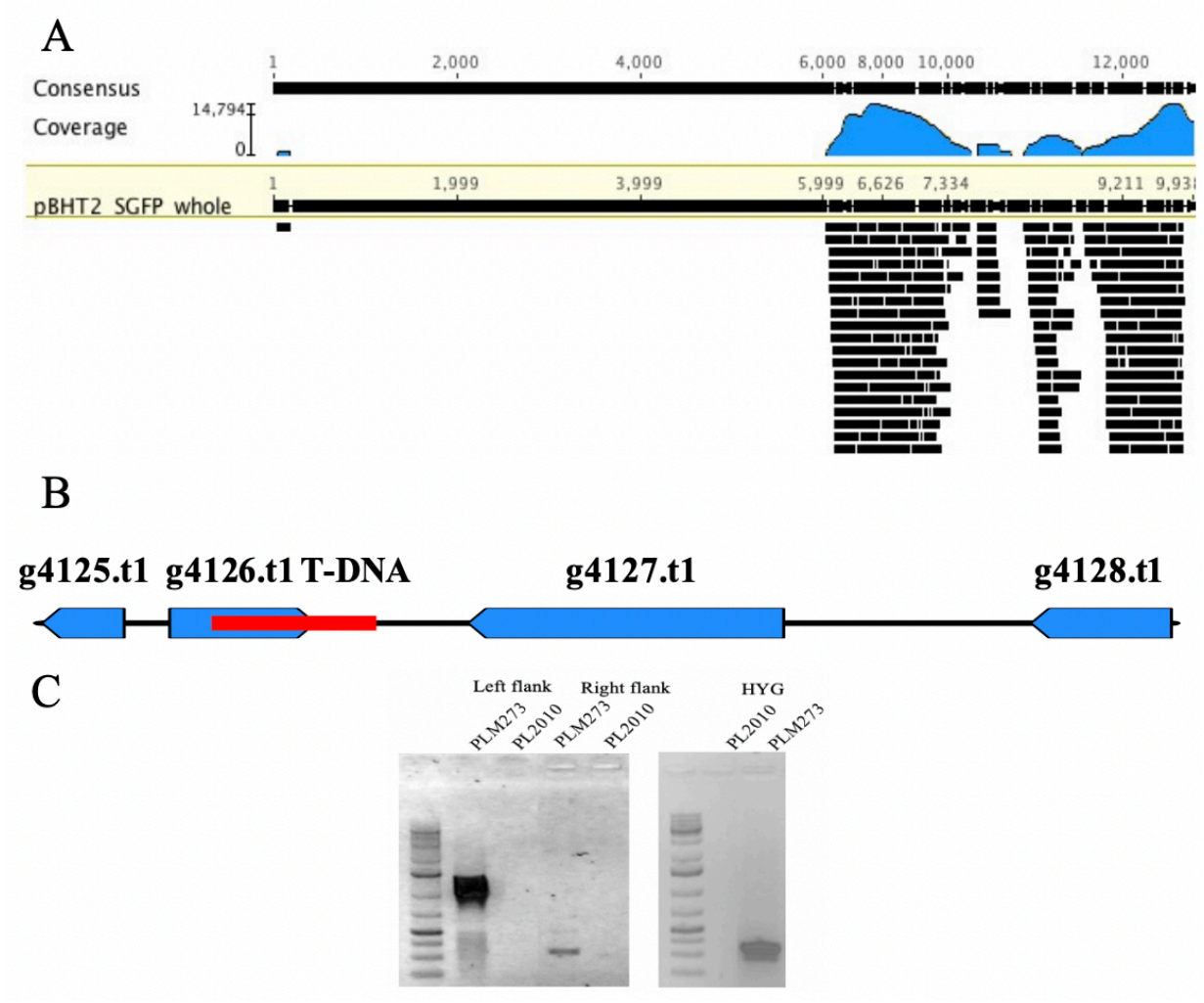


Figure 13. The location of T-DNA of the mutant PLM2739. **(A)** Coverage and presence of reads of T-DNA of the mutant, PLM2739 mapped to the plasmid, pBHt2_sGFP. **(B)** The site of T-DNA in the gene g4126.t. **(C)** Bands of overlapping flanking regions of g4126.t1 with the inserted cassette (T-DNA) and a fragment of HYG compared with the wide-type strain, PL2010.

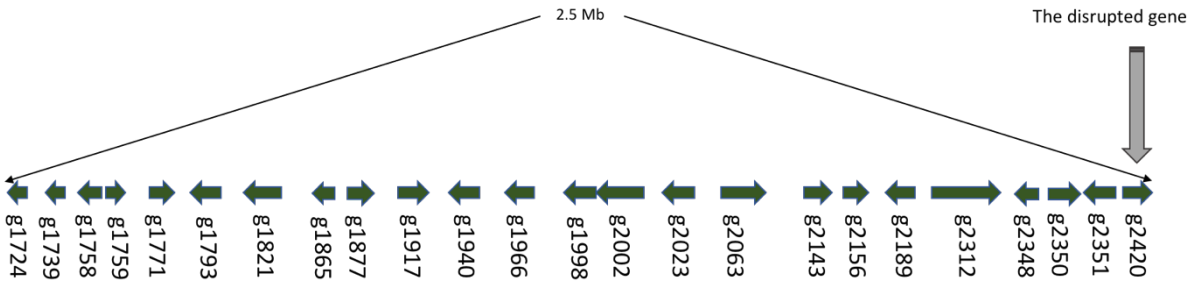


Figure 14. Distribution and order of CYP genes in scaffold 12 of *D. longicolla,* including the disrupted gene g2420.t1.

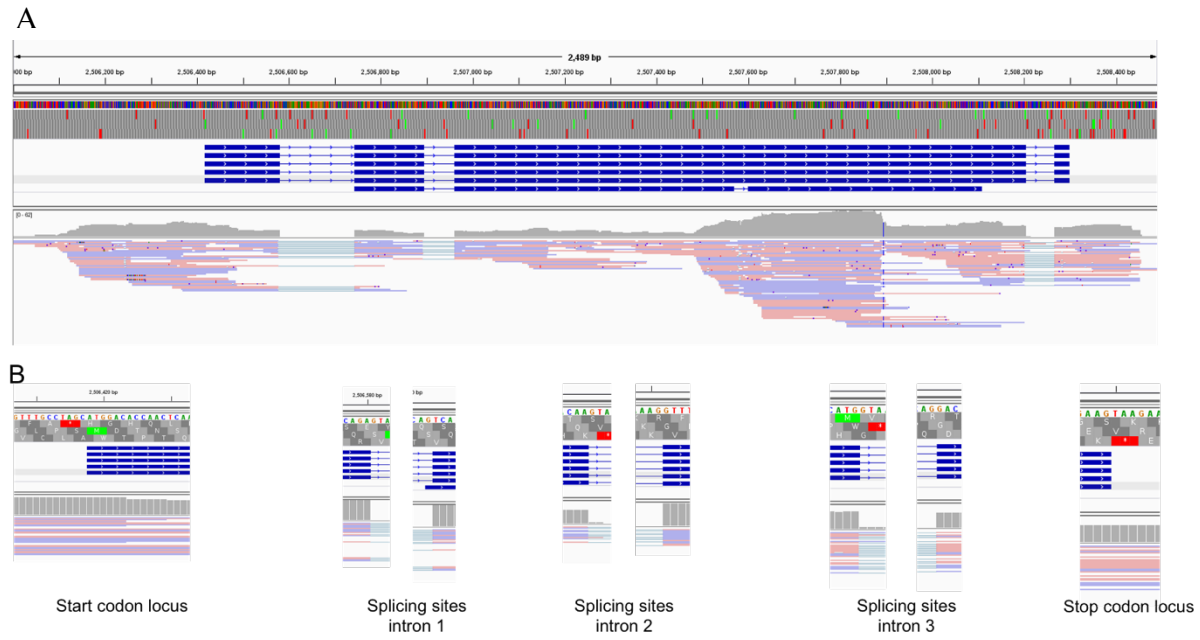


Figure 15. Verification of *D. longicolla* gene g2420.t1. **(A)** Region of gene g2420.t1. Solid blue tracks represent the alignment of homologous proteins (GenBank accessions POS69387.1; ROW02277.1; ROW09819.1; ROW14554.1; KUI68930.1; P54781). RNA-seq coverage is also shown. **(B)** Close-up of g2420.t1 includes the start codon, splicing sites, and the stop codon.

Figure 16. Putative tertiary structure of the CYPX domain of g2420.t1 (cytochrome p450) in *D. longicolla* predicted by the I-TASSER server (Yang & Zhang 2015; Zhang et al., 2017).

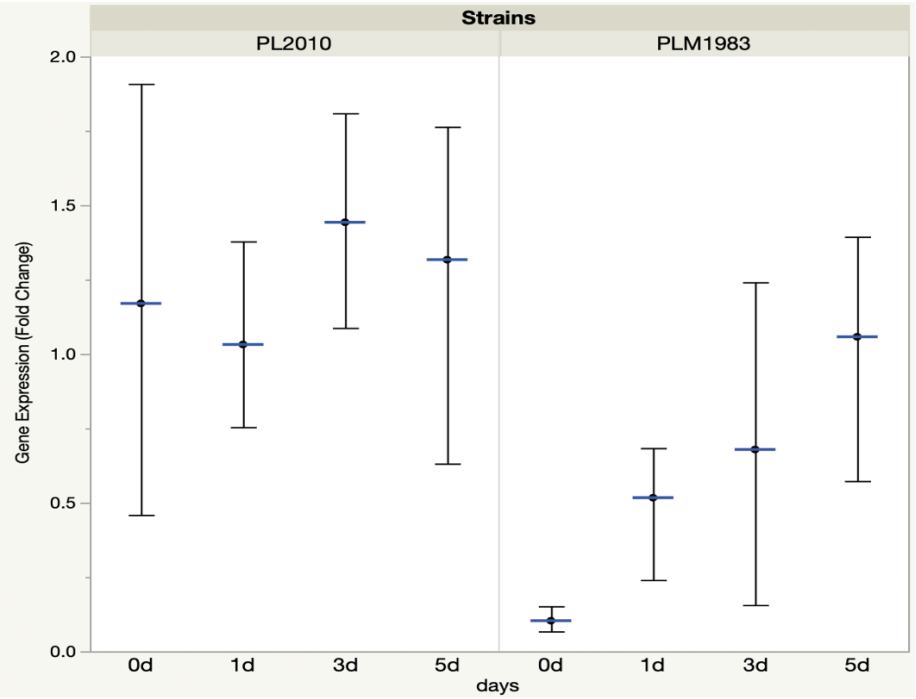


Figure 17. Normalized fold expression (fold change; 2^ΔΔCq) of g2420.t1 (cytochrome p450) in wild-type (PL2010) and the mutant (PLM1983) of *Diaporthe longicolla* at 0, 1, 3, and 5 days after inoculation. The means were calculated from three biological replicates. Bars represent the range of gene expression.

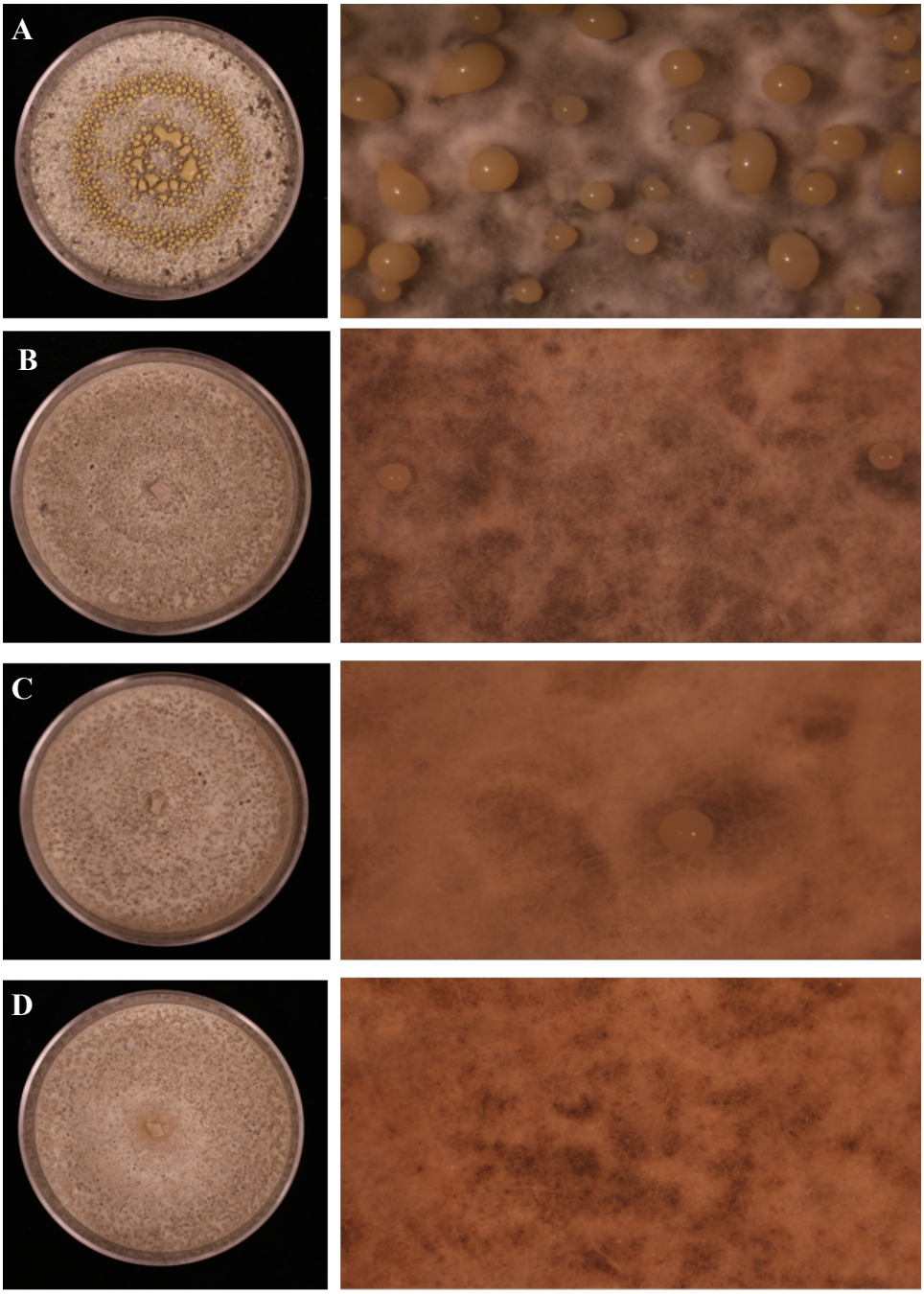


Figure 18. Sporulation of *D. longicolla* strains on PDA after two weeks. **(A)** The wild type strain, PL2010. **(C** and **B)** Complemented strains, comp1983-22 and comp1983-23. **(D)** Mutant strain PLM1983.

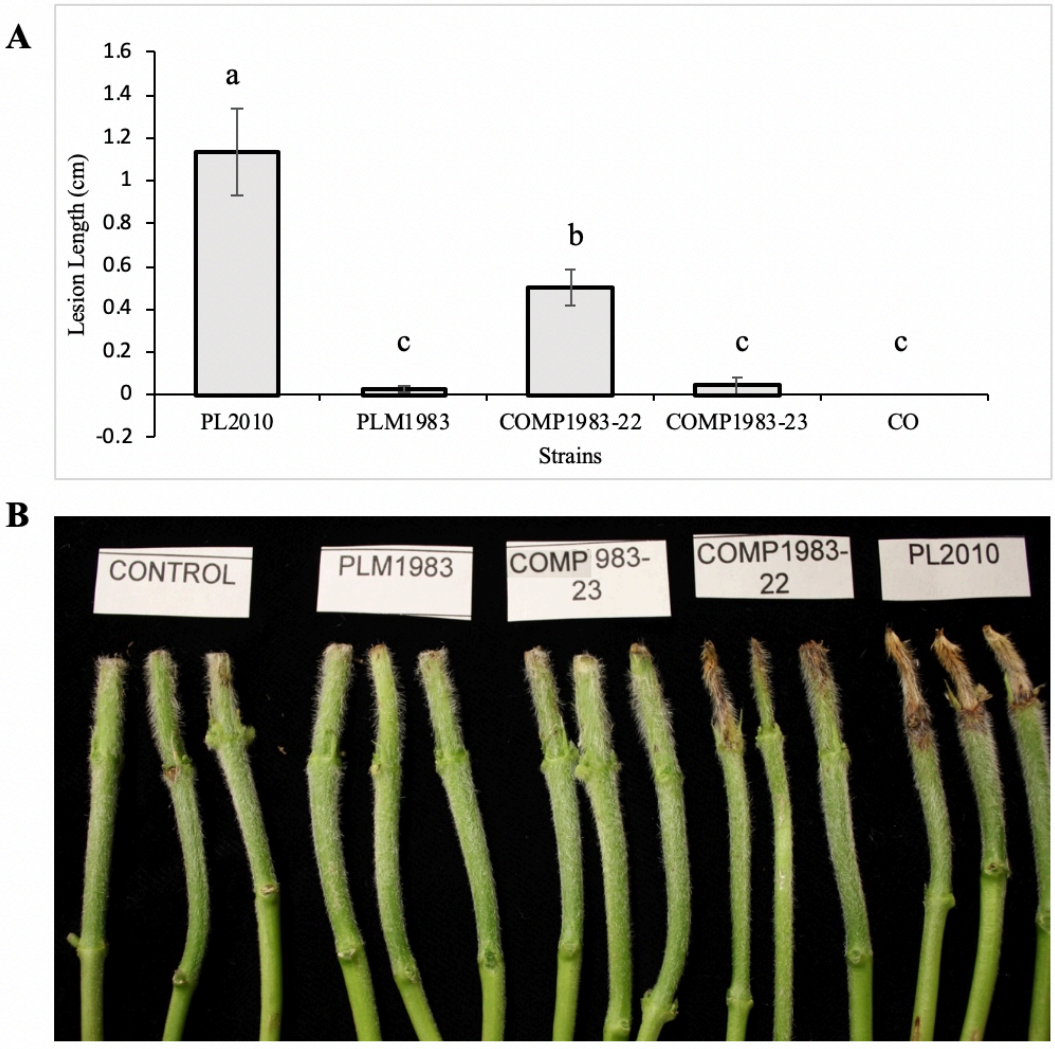


Figure 19. Necrotic lesions on soybean stems caused by mutants (PLM1983), complemented transformants (COMP1983-22 and COMP1983-23) and wild-type (PL2010) of *D. longicolla* with a cut-stem inoculation assay. **(A)** Bars with different letters are significantly different according to the Tukey test ($p \leq 0.05$). **(B)** Necrotic lesions on soybean stems.

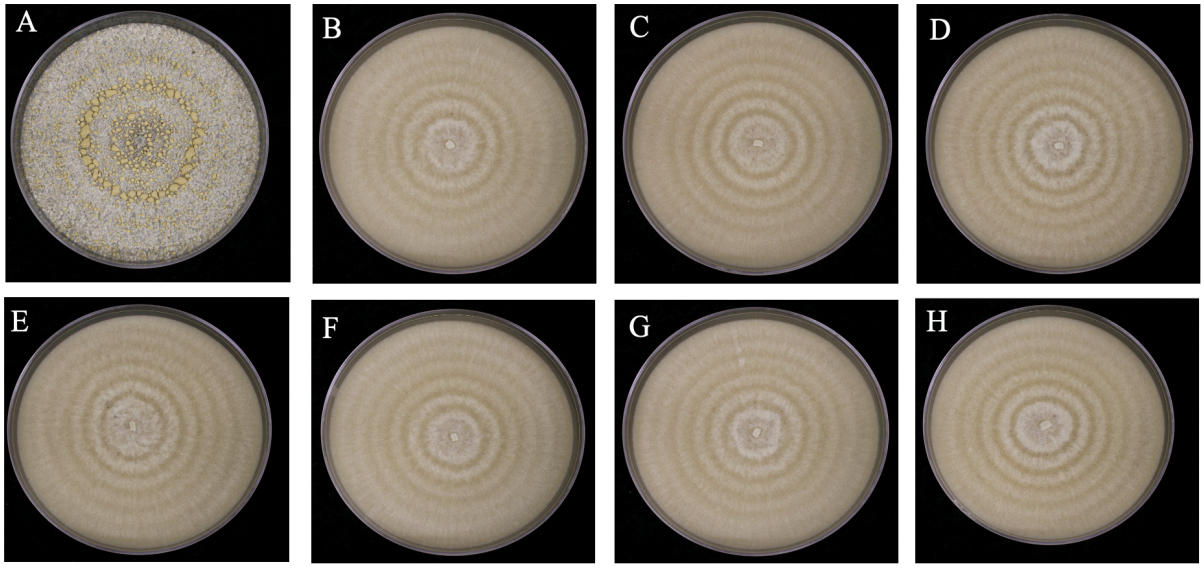


Figure 20. Phenotypes of *D. longicolla* strains on oatmeal agar after two weeks. **(A)** Wild type, PL2010. **(B-C,** and **F- H)** selected transformants of PLM2739 (complementation attempts). **(E)** Mutant strain PLM2739.

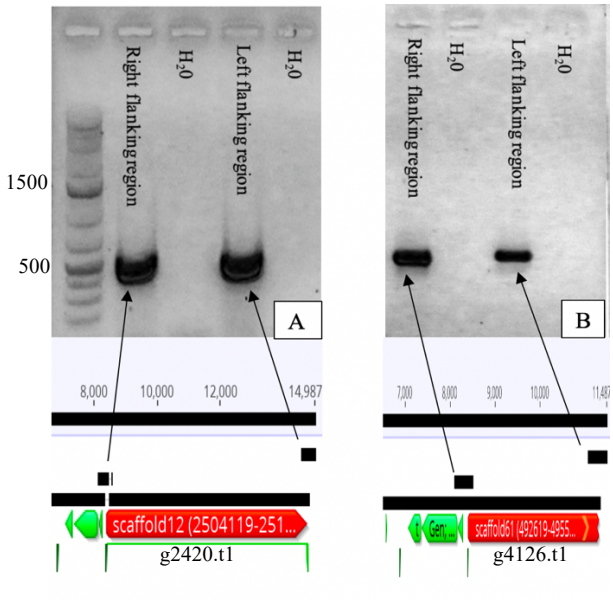


Figure 21. Confirmed reintroduction of g2420.t1 into strain PLM1983 and g4126.t1 into strain PLM2739 via PCR and sequencing. **(A)** Bands of the fragments of flanking regions of g2420.t1 and sequences of these fragments aligned with the cassette of cloned gene. **(B)** Bands of the fragments of flanking regions of g4126.t1 and sequence of the fragments aligned with the cassette of cloned gene.

Table 1. Primers used in this chapter.

Primers	Sequences
g2420qtrexon3-4F	TACGCCATGTCTAACCTC
g2420qtrexon3-4R	AAAGTCAAGGGACAATCG
PlTubF1	CGACAGCAATGGCGTTTACAAC
PlTubR1	ATGGTACCGGGCTCGAGAT
Scf61-TDNA-F1	AGTCAGGACACAAGATCG
Scf61-TDNA-R1	CACCAGATTGAGGTCTCTT
Scf12-TDNA-F1	AGATAGCGGACACCTCTG
Scf12-TDNA-R1	CACTGCACCACGCTTATT
Scf12-TDNA-R2	CGGTCATGGTCTATCATATC
HYGF	CCAGTGATACACATGGGGATCAGC
HYGR	GGATATGTCCTGCGGGTAAATAGCTG
GFPF	ATCCTGGGGCACAAGCTG
G2420compl-F2	gtcgtttttcaactagtgGTCGGGATACGCCTTTG
G2420compl-R2	agctggtgacctcaagctGCAACTTTGAGGGAGCTA
G4126compl-F3	gtcgtttttcaactagtgCTCGCGTGGGAATTCGAAC
G4126compl-R3	agctggtgacctcaagctTGGAAGGATGAGAAGTGGGTC
g2420-right-f	CTCCTCTCAGACAACAC
scf12-right-R	ATTGCGGGACTCTAATC
scf12-R	CACAACCACTTCATTACAC
g4126-right-f	GATGTGATAACATGCAGAA
g4126-right-R	AACCCATCTCATAAATAACG
scf61-R	GATCTTCACGACTAGCAC
gen418-F	CCAGTCATAGCCGAATAG

Chapter 5: Conclusion

The research presented in this dissertation identified three known close morphologically and phylogenetically related *Diaporthe* species, which are *D. longicolla*, *D. unshiuensis,* and *D. ueckerae.* These species are likely a *Diaporthe* species complex since there is a difficulty to identify them using morphological features. Identifying and confirming pathogenicity of *D. unshiuensis* with a relative frequency reaching about 40% in some sites should pay more attention by plant breeders and plant pathologists since this species is not known before on soybean and in the U.S. Additionally, it is morphologically and phylogenetically the closest species to the ubiquitous *Diaporthe* species, *D. longicolla*.

Confirming pathogenicity of endophytic isolates of *Diaporthe* species indicates that these species could have more than one lifestyle and behave as asymptomatic endophytes during the appropriate life stage (s) of soybean beside the necrotrophic lifestyle. Thus, propagules of endophytic *Diaporthe* species during asymptomatic association with soybean could play as an inoculation source for late-season diseases such as Phomopsis seed decay and consequently cause disease's symptoms. More efforts are required to determine transmission ways of *Diaporthe* species from germinated seed through endophytic association until seed colonization and assessing the quantitative correlation between endophytic fungal propagules and disease's severity to reduce the inoculation and manage this disease.

Demonstration of high genetic variability and diversity of *D. longicolla* and *D. unshiuensis* populations in Arkansas by GBS of microsatellites concludes that these species may have high abilities to emerge a genotype with phenotypic traits for surviving and epidemic disease. Since *D. longicolla* are more likely to have clonal populations (asexual populations) while *D. unshiuensis* could have sexual populations according to findings of linkage

disequilibrium, the sources of genetic variation of *D. longicolla* without the probability of sexual recombination still uncertain. However, vegetative compatibility among studied isolates of *D. longicolla* could indicate the possibility of gene flow and parasexual recombination among individuals of this species in geographical sites. Furthermore, no significant differences of genetic variations between populations of every single species across geographic sites and genotypes of each species being not grouped based on geographic sites could approve that pathogen transport and gene flow could occur between geographic sites. Since *Diaporthe* species could be associated with symptomless soybean seeds, this transmission way may play a key role in pathogen transport between Arkansas sites.

The current study identified and characterized the gene cytochrome P450, g2420.t1, potentially involved in seed colonization, the virulence, and the development of asexual reproduction using forward genetic screening. It is suggested that this gene involves the production of secondary metabolites and necrotrophic lifestyle. Additionally, the high number of cytochrome P450 genes in the whole genome of *D. longicolla* encourages further investigation to identify a gene cluster functionally participating with this interesting cytochrome P450.